Title

Allocation of speech in conversation

Journal

Journal of the Experimental Analysis of Behavior, 107(2)

ISSN

0022-5002

Authors

Simon, Carsta

Baum, William M

Publication Date

2017-03-01

DOI

10.1002/jeab.249

JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 2017, **107,** 258–278 NUMBER 2 (MARCH)

ALLOCATION OF SPEECH IN CONVERSATION

CARSTA SIMON[1] AND WILLIAM M. BAUM[2]

[1]OSLO AND AKERSHUS UNIVERSITY COLLEGE OF APPLIED SCIENCES
[2]UNIVERSITY OF CALIFORNIA, DAVIS

In a replication and extension of Conger and Killeen's (1974) widely cited demonstration of matching in conversations, we evaluated nine participants' allocation of speech and gaze to two conversational partners. German speakers participated in two 90-min sessions in which confederates uttered approval on independent variable-interval schedules. In one of the sessions, confederates uttered approval contingent upon and contiguous with eye contact whereas in the other session approval was uttered independent of the participant's gaze. Several measures of participants' verbal behavior were taken, including relative duration and rate of speech and gaze. These were compared to confederates' relative rate of approval and relative duration and rate of talk. The generalized matching equation was fitted to the various relations between participants' behavior and confederates' behavior. Conger and Killeen's results were not replicated; participants' response allocation did not show a systematic relation to the confederates' relative rate of approval. The strongest relations were to overall talk, rather than approval. In both conditions, the participant talked more to the confederate who talked less—inverse or antimatching. Participants' gaze showed the same inverse relation to the confederates' talk. Requiring gaze to be directed toward a confederate for delivery of approval made no difference in the results. The absence of a difference combined with prior research suggests that matching or antimatching in conversations is more likely due to induction than to reinforcement.

Key words: matching law, verbal behavior, contiguity, contingency, conversations, allocation of talk and gaze, humans

In his 1957 publication *Verbal Behavior,* Skinner was the first to outline a natural scientific account of language and speech. Skinner chose the term *verbal behavior* to emphasize the similarities, instead of the differences, between speaking and any other operant behavior. Skinner chose the label *verbal behavior* to highlight his approach to analyzing speaking and other operant behavior by the same means.

Since Skinner's introduction of *Verbal Behavior* in 1957, most behavioral research on people's verbal interactions has focused on enhancement of applied procedures for improving skills of people with communication impairments, as in individuals diagnosed with autism or developmental disorders. The study of verbal behavior has proven fruitful for the applied sector (Schlinger, 2008). However, because Skinner derived his framework from his basic experimental work, the question arises why *Verbal Behavior* has generated mainly applied and little basic experimental research.

One of the reasons may lie in Skinner's choice of unit of analysis. The measure of behavior is usually selected based on the conditions of reinforcement (Baum & Rachlin, 1969). A food delivery contingent upon depression of a lever by 5 mm produces a number of equally short responses that can be counted with an attached microswitch (Baum, 1976). Skinner's framework of how to make sense of humans' verbal interchanges originates from his research on pigeons' key pecking and rats' lever pressing. In those experiments, Skinner measured the target behavior by computing the rate of lever presses and key pecks by dividing their number by the duration of the experimental session. Skinner treated the lever press and the key peck—which are commonly recorded with a pulse former—as discrete events, that is, events that take a considerably shorter time to occur than the time that elapses between those events. This abstraction of discrete responses, which has been the most common way to study activities, is practical when dealing with responses whose variability in duration is negligible. If one wants to analyze a rat's water licking, one can count the licks that occur in one burst, and one can count the

Special thanks go to Professor Per Holth for his assistance in software development. We also thank Prof. Dr. Mark Vollrath for generously providing us with a room without which this study could not have taken place.

Correspondence concerning this article can be addressed to: Carsta Simon, Oslo and Akershus University College of Applied Sciences, Faculty of Health Sciences, Dept. of Behavioral Science, PO Box 4 St. Olavs plass, NO-0130 Oslo, Norway. E-mail: carsta.simon@hioa.no.

doi: 10.1002/jeab.249

bursts of licking per hour. However, other kinds of activities cannot be as straightforwardly analyzed by splitting them into units with negligible duration.

Humans' talking occurs as a stream whose functional units vary greatly in duration. To expand our understanding of verbal behavior, what could an alternative to Skinner's search for tacts, mands, echoics and so forth, look like? The present study's approach to verbal behavior originates in research on choice. All behavior entails choice (Baum, 2013) in the sense that at every moment of its life, an organism engages in one activity or an alternative activity and that over any significant span of time, an organism engages in multiple activities. The observation that behavior is a function not only of reinforcers contingent on that behavior, but also of reinforcers contingent on other activities, carries the potential to uncover lawful relations in people's communicative exchanges in conversations.

Given two concurrently available response alternatives, behavior is distributed approximately in the same proportion that reinforcers are distributed across those alternatives (Baum, 1979). This frequent observation of a linear relationship between relative response rate and relative reinforcer rate is known as the matching relation (Herrnstein, 1961). Herrnstein expressed the relation:

$$\left(\frac{B_R}{B_R + B_L}\right) = \left(\frac{r_R}{r_R + r_L}\right) \quad (1)$$

where B_R and B_L represent the rate of responding to the right and left response alternatives, respectively, and r_R and r_L represent the reinforcer rates for responses to the right and left alternatives, respectively. Based on observed deviations from this strict matching equation, Baum (1974) proposed the generalized matching equation:

$$Log\left(\frac{B_R}{B_L}\right) = sLog\left(\frac{r_R}{r_L}\right) + Log(b) \quad (2)$$

where B_R, B_L, r_R and r_L stand for the same features as in Equation 1.

In the generalized matching equation, the fitted parameters s (the slope) and b (the intercept) represent the behavior's sensitivity to consequences (s) and its bias (b). In perfect matching, $s = 1$ and $b = 1$ (i.e., Log(b) = 0). Sensitivity describes how responsive choice is to differences in the reinforcer ratios. More indifferent responding than the matching law would predict is termed *undermatching*. Such allocation of less behavior to the richer schedule and more behavior to the leaner schedule renders a slope less than one ($s < 1$). A slight deviation from matching ($s = 0.8$) in the direction of indifference is the typical outcome of a concurrent schedule experiment. A more extreme preference than the matching law predicts is called *overmatching*. Overmatching is characterized by slopes greater than one ($s > 1$) and occurs when the participant allocates less behavior to the lean schedule and more behavior to the rich than predicted. Bias not equal to 1.0 indicates preference for one of the response alternatives due to some factor other than the reinforcer ratios.

Unlike lever pressing and key pecking, many kinds of activities, such as reading a book or verbal exchanges, are poorly represented by response counts. The matching relation has also been shown to hold when time spent responding, instead of response rate, is measured (Baum & Rachlin, 1969). For time, the matching equation is written:

$$Log\left(\frac{T_R}{T_L}\right) = sLog\left(\frac{r_R}{r_L}\right) + Log(b) \quad (3)$$

where T_R and T_L represent the time spent responding on the right and left alternatives, respectively. The remaining parameters in Equation 3, describing time allocation, resemble those specified in Equation 2.

Nonverbal response allocation in choice situations has been subject to appreciable basic and applied behavior-analytic research (e.g., Aparicio & Baum, 2009; Baum & Davison, 2014; Cuvo, Lerch, Leurquin, Gaffaney, & Poppen, 1998; Davison & Baum, 2000; Fisher & Mazur, 1997; Schofield & Davison, 1997) but few studies have focused on human verbal interactions. Those studies that have attempted to clarify whether matching theory holds for humans' verbal interactions in conversations have yielded contradictory outcomes, and the question remains to be answered.

Conger and Killeen (1974) were the first to apply time-allocation matching to human verbal interchanges in conversations. They exposed individual participants to a group discussion with

two confederates who pretended to be other participants. The confederates were signaled on two concurrent variable-interval (VI) schedules when to utter a statement of agreement to what the participant had said. On the VI schedules, statements of agreement were uttered after variable and unpredictable time intervals, following a participant's response (Mazur, 2001). On average, 1 min elapsed between signals for both confederates combined. Discussions lasted for 30 min. Confederates switched schedules after 15 min. Conger and Killeen's pooled data indicate that, as opposed to the first 5 min that the participant was exposed to each schedule, during the last 5 min of each schedule, the proportion of participants' duration of talk to each confederate roughly tracked the proportion of agreement statements obtained from that confederate. However, for four of five participants, only one VI VI pair and its reverse were tested. Because a straight line can be fitted perfectly to any two points, Conger and Killeen's data say little about a possible extrapolation to other concurrent VI VI schedules.

Traditionally, a reinforcer is said to strengthen a response or stimulus–response connection when the reinforcer occurs contiguously with the response (Skinner, 1948; Thorndike, 1911/2000). Accordingly, Herrnstein (1961) presented matching as relative response strength, assuming that matching represents the differential strengthening of response alternatives by different rates of reinforcement. Thus, when testing the applicability of Herrnstein's theory of matching as strengthening to verbal behavior, the event assumed to strengthen the verbal response must be presented in close temporal proximity to the verbal response serving as dependent variable. If one is studying strengthening by reinforcement, logic demands that one measures the activity that is supposed to be reinforced. In Herrnstein's experiment, pecks were followed by food, and pecking was measured. However, Conger and Killeen (1974) did not deliver approval contiguous with their dependent variable. Although they measured time looking at a confederate (gaze) when talking as their dependent variable, approval (the putative reinforcer) was delivered for any verbal behavior—that is, independently of gaze. This lapse opens the possibility that Conger and Killeen's results do not represent strengthening by reinforcement.

Alternatively, the approval of a confederate may have drawn the participant's gaze toward that confederate, and the more approving confederate drew more gaze. This possibility accords with the interpretation that Conger and Killeen's (1974) results represent induction rather than strengthening (Baum, 2012; Segal, 1972). The more approving confederate induced more gaze, and the less approving confederate induced less gaze.

A way to test the possibility that Conger and Killeen's (1974) results may be explained by induction is to introduce a contingency in which approval depends on gaze as well as verbal behavior. If strengthening matters, then when this condition is compared with their original condition, we should see a difference—presumably more effect of approval on gaze when approval is delivered contingent upon gaze. In contrast, if no difference results, we may conclude that Conger and Killeen's results do not represent strengthening and possibly represent induction.

To further examine allocation of verbal responses under concurrent schedules of approval delivery, Borrero et al. (2007) conducted a study similar to Conger and Killeen's (1974). In Borrero et al.'s study, 25 participants conversed individually with two confederates who uttered agreement on a dense and a lean schedule over the course of either 20 or 30 min. Their individual data from seven of nine participants (i.e., those participants who yielded enough data points to be included into analyses) were better described by the response-rate equation (Eq. 2; counting utterances) than by the time-allocation equation (Eq. 3). However, this analysis including all time intervals does not allow comparison of the results to other studies in which goodness of fit was evaluated during the first and last 5-min intervals of each session. For a comparison of those intervals, Borrero et al. pooled all participants' data. Contrary to Conger and Killeen's pooled results, their pooled data from the first 5 min were better described by the matching law than the data from the last 5 min of each session.

Pierce, Epling, and Greer (1981) studied verbal matching behavior in longer and more artificial discussion sessions. They exposed each of six participants to seven 1-hr discussion sessions. During those sessions, each participant, who was equipped with a throat microphone, chose

which of two confederates she wanted to speak to. By operating a change switch, the participant could freely choose to whom to talk but could only talk to one confederate at a time. The selected confederate agreed with the participant's statements according to seven different predetermined VI schedules. The participants' relative frequency and duration of talking were fitted to Equations 2 and 3. Pierce et al.'s results showed high variability but generally, slopes (*s*) were higher during the second half hour, after the participants had had experience with the schedule in effect. However, only two participants' relative rate or relative duration of talking produced positive slopes, indicating severe undermatching. The remaining five participants' relative durations and rates of responses were either indifferent between response alternatives or produced negative slopes, which mean that the participant talked longer and more frequently to the confederate who expressed fewer statements of agreement.

McDowell and Caron (2010) published an analysis of 20-min natural conversations of 210 pairs of boys consisting of one boy at risk for juvenile delinquency and his peer. The boys' utterances were classified either as positive or nonpositive social responses, as normative talk, or as rule-break talk. A positive social response such as an approving statement or a smile was assumed to reinforce the other boy's verbal bout if it directly followed it. Nonpositive responses included all remaining responses together with being silent. Statements containing a violation of a legal or conventional norm (such as lying to parents) were classified as rule-break talk, which constituted response Alternative 1. Normative statements including, for example, talk about friends, family or school, were regarded as response Alternative 2.

McDowell and Caron (2010) fitted Equations 2 and 3 to each boy's frequency and duration of talking bouts and found matching to describe the boys' verbal behavior accurately. Besides the dissimilarity in target behavior and consequences (approval by two other conversational partners vs. positive and nonpositive responses by one peer), McDowell and Caron's study differed from earlier research in that they did not program the putative reinforcers. At first sight, this use of a natural setting may seem to render their findings more externally valid than those of studies in laboratory settings. However, because they did not manipulate the allocation of positive responses, the latter cannot be interpreted as an independent variable. From the reported observations, one cannot tell whether one boy's positive and nonpositive statements caused the other boy's allocation of talking bouts to rule-break and normative talk or the other way around. McDowell and Caron's observations might not illuminate the cause–effect relation but could still shed light on how two conversational partners' utterances affect each other in some kind of feedback loop[1].

Taken together, the studies by Conger and Killeen (1974), Pierce et al. (1981), Borrero et al. (2007), and McDowell and Caron (2010) leave the matter of verbal matching behavior unresolved. The present study was designed to specify the conditions under which verbal stimuli in conversations control response and time allocation. The earlier studies examined the effect of approval statements on the participants' duration or frequency of talk. We also examined other

[1]Moreover, delivery of the positive and nonpositive consequences in McDowell and Caron's (2010) study may not have resembled a concurrent VI VI schedule. On the contrary, allocation of statements probably occurred on two equal ratio-like schedules. If so, response allocation could not vary from reinforcer allocation. On ratio schedules, reinforcer delivery depends solely on the number of responses executed. In ratio schedules, we may define the probability (p_i) of reinforcer (r_i) per response (B_i) as

$$p_i = r_i/B_i \quad (1)$$

If applied to two VR schedules approximating equal probabilities of reinforcer delivery ($p1 \approx p2$), Equation 4 becomes

$$r_1/B_1 = r_2/B_2 \quad (2)$$

In McDowell and Caron's study, B1 represents rule-break talk, B2 stands for normative statements, r1 represents a positive response, and r2 represents a nonpositive response—putative reinforcers differing in magnitude. If we rearrange Equation 5, we arrive at

$$B_1/B_2 = r_1/r_2 \quad (3)$$

the matching equation. Equation 6 does not describe the outcome of a behavioral process but environmental contingencies that are inevitably true for concurrent VR VR schedules with equal probabilities of reinforcer delivery. Any distribution of responses obeys the matching law when $p1 \approx p2$ (Herrnstein & Vaughan, 1980). Since conversations lasted for 20 min only, even if unequal VR schedules were in effect, matching might have been forced.

variables that can affect the behavior of conversational partners, modelled by confederates and participants. The following verbal stimuli produced by the confederates were examined: a) approval statements, b) the duration of speech directed at the participant, c) the frequency of speech directed at the participant, and d) the duration of overall speech directed at the participant.

The present study attempted to replicate and extend the work of Conger and Killeen (1974) to assess the effects of contiguity of approval on participants' verbal behavior and to explore additional characteristics of two confederates' verbal behavior that might affect a participants' speech. The specific objectives were to investigate whether participants match their relative rates and durations of verbal responses to the relative rates of verbal approval when directing comments to two confederates, whether the relative duration and frequency of confederates' overall speech directed to the participant affects the participant's allocation of talk and gaze between the two confederates, whether participants' allocation of verbal statements and eye-gazes is better described by Equation 2 (relative response rates) or Equation 3 (relative time allocation), and whether allocation of verbal responses and eye-gazes to two different confederates differs when the confederates' approval is contiguous with gaze (condition GR—gaze required) versus when approval need not be contiguous with gaze (condition GN—gaze not required).

Method

Participants

Nine native German speakers participated in the study. Five of them were female. They had a median age of 24 years, ranging from 20 to 66. Participants had diverse educational backgrounds; all had completed high school. Five of the participants were currently graduate or undergraduate students of psychology at the Braunschweig University of Technology, Germany. Four participants received course credit for their participation. Those course credits could also have been earned in alternative ways. The remaining participants, who did not need course credits, participated "out of interest" and did not receive any explicit reward. Participants were recruited via a psychology student mailing list and through word-of-mouth among one author's acquaintances. They were told that the study examined people's responses in a videotaped three-participant discussion group on topics of interest to everyone. Each participant attended two 90-min sessions, scheduled on two consecutive days for seven of the participants and two days apart for two of the participants. Applicants who were acquainted with either of the confederates were excluded from participation. The study was performed in adherence with the Helsinki Declaration and was approved by the relevant local ethics board.

Confederates were two female psychology undergraduates aged 20 and 22 who had similar first names and looked and dressed very much alike. The moderator was a male psychology undergraduate aged 26. The same confederates and moderator participated in all sessions.

Setting and Apparatus

Prior to the first experimental session, the moderator asked the participant and ostensibly the confederates to sign an informed consent form and a topic choice form in a room next to the experimental room. Participants were told that they would participate in a discussion to assess the development of everyday life topics in conversations. The moderator explained further that a total of six topics would be discussed in the course of the two 90-min sessions that the participant was asked to participate in. The consent form also reminded them that the discussion would be videotaped and asked participants not to talk about the experiment for another 2 months during which data were collected. It also explained that it was possible to collect course credit, that participation was voluntary, and that data would be anonymized and treated confidentially.

The topic choice form consisted of a list of 13 topics, each described in about three sentences. Examples of topics were school fees, Facebook, student leagues, polyamory, the connection between violence and computer games, and volunteer work. The participant and, for appearances, the confederates were instructed to cross out five topics that they would by no means want to discuss and to rank the remaining topics by preference.

The topics actually discussed were chosen from the participant's list.

After the forms were signed, the moderator directed the participant and the two confederates into the experimental room, which measured approximately 8 m x 4 m. Materials included a table measuring 80 cm x 80 cm, four chairs (one for the participant, one for each confederate, and one for the moderator responsible for evoking discussion), and a carafe with water and four glasses placed in the middle of the table. The room also contained a general clutter consisting of bookshelves, desks, chairs, and PCs, which were expected to mask the presence of a monitor showing lights which signaled the confederates when to utter approval.

One participant at a time was involved in a discussion session. The participant was strategically seated with his or her back to the monitor that signaled the confederates when approval statements were due. The screen was partly hidden behind a wall projection. The confederates were seated to the participant's left and right; the moderator sat directly across from the participant. Sessions were videotaped using a clearly visible digital camera mounted on a tripod behind the moderator, facing the participant and recording the whole group. Green and yellow lights for signaling times at which the confederates were to express approval appeared on the monitor behind the participant's head. The occurrence of the lights was programmed in Visual Basic. Video recordings were viewed for coding using ELAN[2] (Eudico Language Annotator).

At the end of the second session, participants filled out a questionnaire that probed suspicion of the confederates' roles, collected demographic information, and asked who of the seemingly other two participants they would prefer to talk to again if there were an additional experiment and why.

[2]ELAN is a professional open-source tool for the creation of complex annotations on video and audio resources. It is provided by the Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, The Netherlands: http://tla.mpi.nl/tools/tla-tools/elan/ (Sloetjes & Wittenburg, 2008; Wittenburg, Brugman, Russel, Klassmann, & Sloetjes, 2006).

Procedure

Participants were asked to wait in a hallway prior to the sessions. Confederates waited there also and pretended to be unacquainted. Before the start of the first session, the moderator gave the consent form and the topic choice form described above to the participant and the two confederates. All topic forms showed a number in the upper right hand corner. The confederates' topic forms contained the same numbers during all sessions; the participants' forms each contained an individual consecutive number. The moderator pointed out that those numbers were also depicted on signs placed on the table around which the discussants would be seated after delivering the topic choice forms. This procedure guaranteed that seating was controlled but appeared to be random to the participant. The seating assigned by those place cards remained constant across sessions.

All participants discussed those six topics that they ranked highest. The order of topics was random with respect to the order in which the participant had ranked the topics. Confederates were unaware of the participant's topic preferences. During the sessions, the participant sat at one end of the table directly across from the moderator and the video camera. The confederates sat on either side. After everyone was seated, the moderator started the camera, read the short description of the topic and started the program that operated the lights to signal the confederates when to deliver statements of approval. The program was started by a click on a wireless mouse that appeared to belong to the laptop from which the moderator had read the topic descriptions.

The session began when the moderator announced the topic. He encouraged the group to begin talking and mentioned that it was not a problem if they got sidetracked. After that, the moderator spoke only to ask questions prepared beforehand when the conversation lagged. He kept his eyes focused on the laptop in front of him and only looked up when he was speaking. Each confederate provided verbal agreement when a designated light flashed and the participant was uttering statements that one could approve of. In condition GN (gaze not required), the confederate's utterance of agreement statements was independent of the participant's gaze. In

condition GR (gaze required), a confederate only expressed approval when the participant looked in her direction so that she could catch the participant's eye and maintain eye contact while making her statement. If a light flashed when a participant was not speaking, the approval statement was delivered the next time the participant spoke, as long as the other confederate's light had not flashed in the meantime. Due to these lost opportunities, obtained reinforcer rates often differed from the programmed rates. Because of the additional contingency in condition GR, fewer approval statements could be uttered in that condition. Five of the participants were first exposed to condition GN, and the remaining four participants started with condition GR.

Confederates were instructed to talk as though in normal conversation, with these restrictions: 1) keep eyes moving while talking, look at each other, the moderator or somewhere else in the room when not talking and do not focus on the participant for more than 2 s unless uttering approval after a signal; 2) if the other confederate makes a statement, then make a statement of similar length; and 3) keep opinions similar to the other confederate and neutral or moderate. Verbal agreement was defined as looking at the participant while making any statement of agreement, for example, "Yes, I agree", "Exactly", or "Good point". Confederates were instructed to refrain from nodding and smiling at all other times.

Scheduled overall reinforcer rate was held constant at 15 per 5-min interval. When each interval timed out, one or the other light on the screen behind the participant was activated according to a probability that was uninfluenced by the participant's or confederates' behavior. Interlight intervals ranged from 5 to 30 s. The programmed distribution of VI schedule intervals was arranged by picking a series of quasirandomized digits between 5 and 30, with a mean of 20, and with a higher number of digits toward the lower end of the scale, roughly according to the Fleshler-Hoffman distribution (Fleshler & Hoffman, 1962). In each of the two 90-min sessions, participants experienced six concurrent VI VI schedules of approval statements. The scheduled approval ratios, each in effect for 15 min, were 13:2, 2:13, 6:9, 9:6, 4:11, and 11:4. Lights were visible for 5 s. Thirty minutes after the lights corresponding to the first (and third) schedules were illuminated, the moderator announced the switch to the next topic. This way, all participants discussed a total of six topics in two 90-min sessions on 2 separate days. One of these sessions occurred under condition GN, one under condition GR. In each of the sessions, participants were exposed to all schedules in the order given above.

After terminating the first session, the moderator asked if any of the discussants had any comments (which no one had) and reminded the group of the second meeting. After that, confederates left the building together with the participant and walked off in opposite directions.

At the end of the second session, the moderator thanked the participants for their participation and asked them to fill out the questionnaire that collected demographic information and encouraged participants to write down "any thoughts regarding the experiment". After that, the moderator announced that he would like to talk to "every participant separately" and pretended to choose the participant first. In confidence, he first asked generally if the participant had anything he or she would like to share regarding the experiment. Then, he asked directly if she had noticed that the other two "participants" were confederates and debriefed the participant on the objectives and procedures of the study. No participant had a suspicion. All were surprised and curious. None of the participants seemed uncomfortable, and when asked by the moderator, many reported they understood the necessity of such "common social psychology procedures," and none of them said they were concerned, disappointed, angry or the like. After that, the confederates entered the room, and the participant was offered the opportunity to ask questions. In the pre-debriefing questionnaire and when asked by the moderator, no participant reported noticing the lights, or the use of any equipment other than the video camera.

Coding

All sessions (i.e., the total of 180 min of conversation per participant) were coded in 5-min segments for:

a. the duration that the participant talked continuously (ignoring pauses less than 1 s) to the confederate seated to the right

(Con R) and to the confederate seated to the left (Con L); TT_R and TT_L

b. the duration that the participant looked at Con R and Con L; TG_R and TG_L
c. the number of times per 5 min that the participant started talking to Con R and Con L; RT_R and RT_L, after not having talked for at least 3 s
d. the number of times per 5 min that the participant started looking at Con R and Con L; RG_R and RG_L, after not having gazed at the confederate for at least 3 s
e. the number of times per 5 min that the participant started talking following a confederate's approval statement
f. the number of approval statements per 5 min uttered by Con R and by Con L; ref_R and ref_L
g. the duration that Con R and Con L each talked to the participant and talked otherwise
h. the number of times per 5 min that Con R and Con L each talked to the participant and talked otherwise

Four trained assistants coded 5-min segments. If the participant was silent for a fraction of a second and continued to look at the confederate, timing of participant's talk was not interrupted. Time spans exceeding 1 s of no speech or gaze towards the confederate did not contribute to the participants' duration of talk measure. All coding was done blind to condition and to the exact schedules in use (the lights on the monitor may have given a general idea, however). Coders observed the video recording and simultaneously pressed one of two keys on a laptop computer whenever the target behavior occurred. ELAN tabulated the data for later analysis.

To double-check the implementation of condition GN and GR a sample of 11.4% of the recordings of sessions under both conditions was also coded for the number of times that the participant and a confederate actually had eye contact when the latter uttered her approval statement.

Table 1 lists the operational definitions of the target behaviors.

Interrater Reliability

Of the data on the participant's behavior, 19.4% was either coded twice or three times by different coders to check interrater agreement. Coder reliability proved to be high, with a correlation between measures recorded by first and second (or third) coders ranging from 0.92 to 0.99 (with a mean of 0.97).

Procedural Reliability

To check the enforcement of the eye contact-independent approval statements in condition GN and eye contact-contingent approval in condition GR, the percentage of approval statements delivered with and without eye contact was calculated in a sample of both conditions. In condition GN, 43.7% of approval utterances by the right confederate and 36.3% by the left confederate occurred during eye contact (i.e., 56.3% and 63.7% occurred without eye contact). In condition GR, all approval occurred during eye contact.

Results

Analyses

To examine our measures of choice for conformance to matching we first compared single participants' behavior. Second, we fitted slopes across participants when bias was factored out. Third, we analyzed raw data summed across participants and, fourth, we performed a multilevel analysis. To allow time for participants' behavior to become susceptible to the current schedule of approval, analyses were based on performance during the last 5 min of each schedule. All logarithms were calculated to the base 2.

Verifying the Procedure

To evaluate if obtained relative approval rates varied sufficiently to potentially gain control of behavior, programmed versus obtained approval rates were plotted in Figure 1 for conditions GN and GR. X-axes show programmed log ratios of approval, and y-axes show obtained log approval ratios. Although the slope of 0.7 indicates some compression in the range of obtained log ratios in condition GR, requiring gaze for approval to occur did not affect the delivered approval distribution substantially.

Although the same number of approval statements was scheduled for both conditions

Table 1

Operational definitions of coded target behaviors (dependent and independent variables) per 5-min interval.

Name	Code	Definition
Participant's duration of talk directed at confederate	TT_R, TT_L	Time that the participant talks while looking at (= orientating her pupils toward) the confederate on her right (TT_R) and her left (TT_L)
Participant's rate of talk directed at confederate	RT_R, RT_L	Number of times that the participant starts to talk while looking at (= orientating her pupils toward) the confederate on her right (RT_R) and her left (RT_L), after either a confederate had talked or three seconds of silence had passed
Participant's duration of gaze directed at confederate	TG_R, TG_L	Time that the participant is looking at (i.e., is orientating her pupils toward) the confederate on her right (TG_R) and her left (TG_L)
Participant's rate of gaze directed at confederate	RG_R, RG_L	Number of times per 5-minute interval that the participant starts looking at (i.e., is orientating her pupils toward) the confederate on her right (RG_R) and her left (RG_L) after not having looked at her for at least three seconds
Confederates' number of approval statements	ref_R, ref_L	Number of times the right (ref_R) or left (ref_L) confederate looks at and verbally agrees with the participant (e.g., says "Yes, that's right", "Exactly!") during or directly after the participant's speaking. In both conditions GN and GR those statements were to occur as soon as possible after the designated light flashed and before the other confederate's light was turned on. In condition GR, the participant's pupils were also required to be directed at the confederate that uttered the agreement.
Duration that Con R and Con L talk directed at participant and talk otherwise	t_R, t_L, t_{R_O}, t_{L_O}	Time that the right (t_R) and the left (t_L) confederate talk while having their pupils oriented towards the participant and the remaining time of the right (t_{R_O}) and left (t_{L_o}) confederate's talk
Rate that Con R and Con L talk directed at participant and talk otherwise	r_R, r_L, r_{R_O}, r_{L_O}	Number of times per 5-minute interval that the right (r_R) and the left (r_L) confederate start talking after a three second silence or after the participant had talked, while having their pupils oriented towards the participant and number of the remaining times of the right (r_{R_O}) and left (r_{L_o}) confederate's first words after having not talked for at least three consecutive seconds

only 80.5% of the total number of approval statements given in condition GN were given in condition GR. Although 19.5% less approval was given in total in condition GR, participants on average talked 12% longer and 11% more frequently to both confederates together (five out of nine participants talked longer and more frequently in condition GR). Thus, overall approval uttered by confederates did not induce the participants to talk more in general.

On average, participants started talking to the confederate who had just delivered approval following 61% of approval statements in condition GN and after 63% of approval statements in condition GR. Of the nine participants, four did so more frequently in condition GN and four in condition GR, which indicates that any speech- and gaze-inducing effect of a confederate's approval statement was independent of whether approval was contiguous with gaze or not.

Comparing Slopes within Individual Participants

As a comparison, each panel in Figure 2 shows the regression slope when gaze was required plotted against the regression slope when gaze was not required. If the two slopes were equal for a participant, that participant's point would lie on the major diagonal (dashed

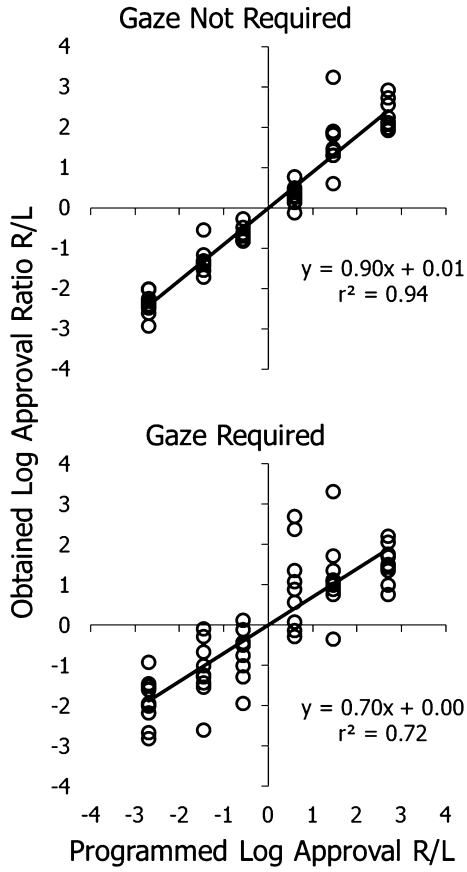


Fig. 1. Obtained versus programmed log approval ratios. X-axes show programmed log ratios of approval and y-axes show obtained log approval ratios. The upper part of the Fig. shows condition GN (Gaze Not Required for approval) and the lower part shows condition GR (Gaze Required for approval).

line). The left two graphs plot slopes of log-ratio duration measures (log TT_R/TT_L, "P-Talk," in the top graph and log TG_R/TG_L, "P-Gaze," in the bottom graph) regressed against log-ratio approval ("C-Approval," log ref_R/ref_L). The scatter of points to both sides of the dashed line indicates that requiring gaze for approval had no systematic effect on slope in comparison with not requiring gaze. The slopes for approval ratio were generally small and positive, with a few outliers in quadrants I and II.

Evidently, the confederates' approval exerted little control over the participants' talk and gaze. If approval per se was unimportant, other aspects of the conversation might have been. In particular, the confederates differed in how long they talked to the participants. Accordingly, we investigated how the difference in confederates' talk might have affected the participants' talk and gaze.

The two graphs on the right side of Figure 2 show the slopes for the log-ratio duration measures regressed against log-ratio of the duration of confederates' duration of talk to the participant. In contrast to the slopes for approval ratio on the left, these graphs indicate that most slopes were negative. When both slopes for a participant were negative, the point for that participant would fall in quadrant III. With one exception—a point in quadrant IV—most points fall clearly in quadrant III or close to the x- or y-axis. For P-Talk ratio (top), 13 out of 18 slopes were negative, and for R-Gaze ratio (bottom), all 18 slopes were negative. A negative slope indicates that the participant spoke relatively more to the confederate who spoke relatively less. For reasons that remain to be understood, participants tended to talk to the confederate who was less talkative in preference to the confederate who was more talkative.

The difference in size and sign of slopes between log-ratio of approval and log-ratio of confederates' talk seems to be all that can be reliably concluded from Figure 2. Few of these slopes were well-determined (see Table A1). When P-Talk was predicted by C-Approval, r^2 ranged from .04 to .94 with a mean of .30 for condition GN and ranged from .01 to .30 with a mean of .11 for condition GR. When P-Talk was predicted by C-Talk, r^2 ranged from .00 to .78 with a mean of .27 for condition GN and ranged from .00 to .83 with a mean of .38 for condition GR.

The results for relative number of speech and gaze episodes were about the same. Least-squares regression was used to fit a line to the log-ratio of rate of talk (RT_R/RT_L) as a function of log-ratio of approval statements (ref_R/ref_L; C-Approval), and between log-ratio of number of looks (RG_R/RG_L) as a function of log-ratio of approval statements (C-Approval). The regression slopes showed the same patterns as in Figure 2, and the proportion of variance accounted for by rate measures was generally low, suggesting caution on any conclusions based on slopes (see Table A2 for r^2s). Whereas Borrero et al. (2007) found that their data were best described by the response-rate matching equation (Eq. 2), fitting our

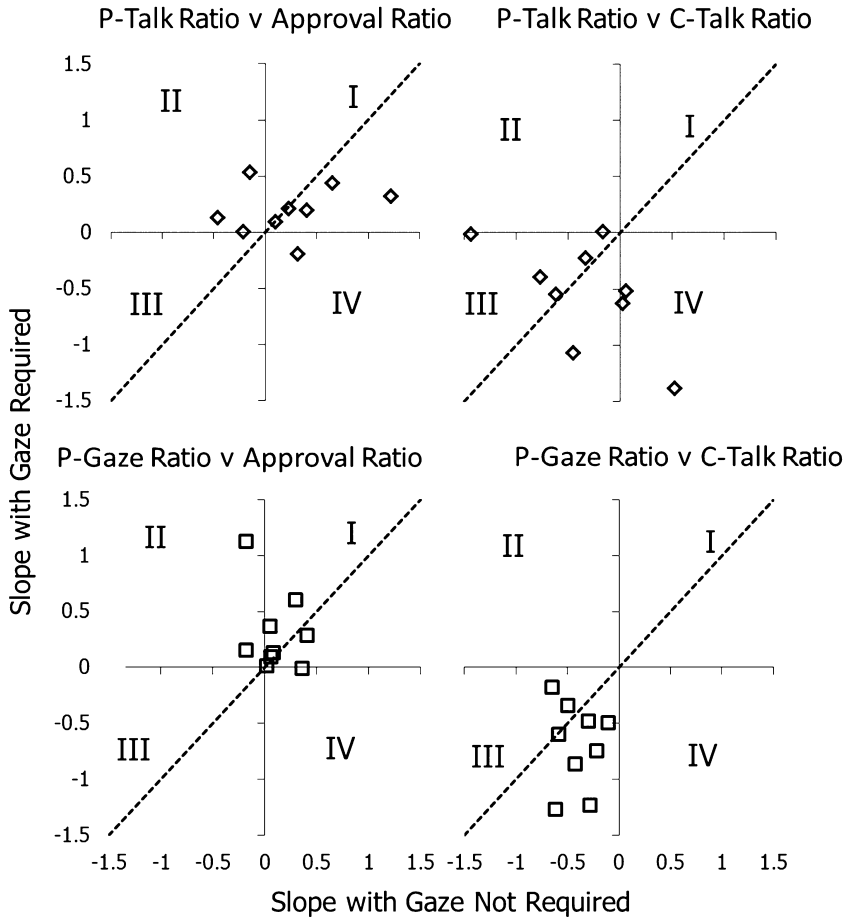


Fig. 2. Individual regression slopes for gaze and talk in conditions GR (Gaze Required for approval) and GN (Gaze Not Required for approval). The left panels show slopes of log-ratio duration measures of talk (upper left) and gaze (lower left) regressed against log-ratio approval. The right panels show slopes of log-ratio duration measures of talk (upper right) and gaze (lower right) regressed against log-ratio confederates' talk. Participants' points lying on the dashed line indicate equal slopes.

gaze and speech data to Equations 2 and 3 showed response duration to be an equally good measure of speech and gaze allocation as response rate.

Choice Pooled Across Participants

Conger and Killeen (1974) plotted pooled results for all their participants. We did the same, but because we had six behavior ratios per approval ratio, we were able to correct for differences in bias across participants, thereby eliminating some of the variability across participants. The correction derived from a rearrangement of Equation 3:

$$Log\left(\frac{T_R}{T_L}\right) - Log(b) = sLog\left(\frac{r_R}{r_L}\right) \qquad (4)$$

The corrected log-ratio measure is the expression on the left of Equation 7. Each panel in Figure 3 shows the corrected log ratio of pooled participants' duration of talk to the left and right confederate plotted against confederates' log ratio of approval (left two panels) and confederates' log ratio of duration of talk (right two panels). Correcting participants' talk log ratios eliminated bias as a source of variability across participants, allowing a better comparison of

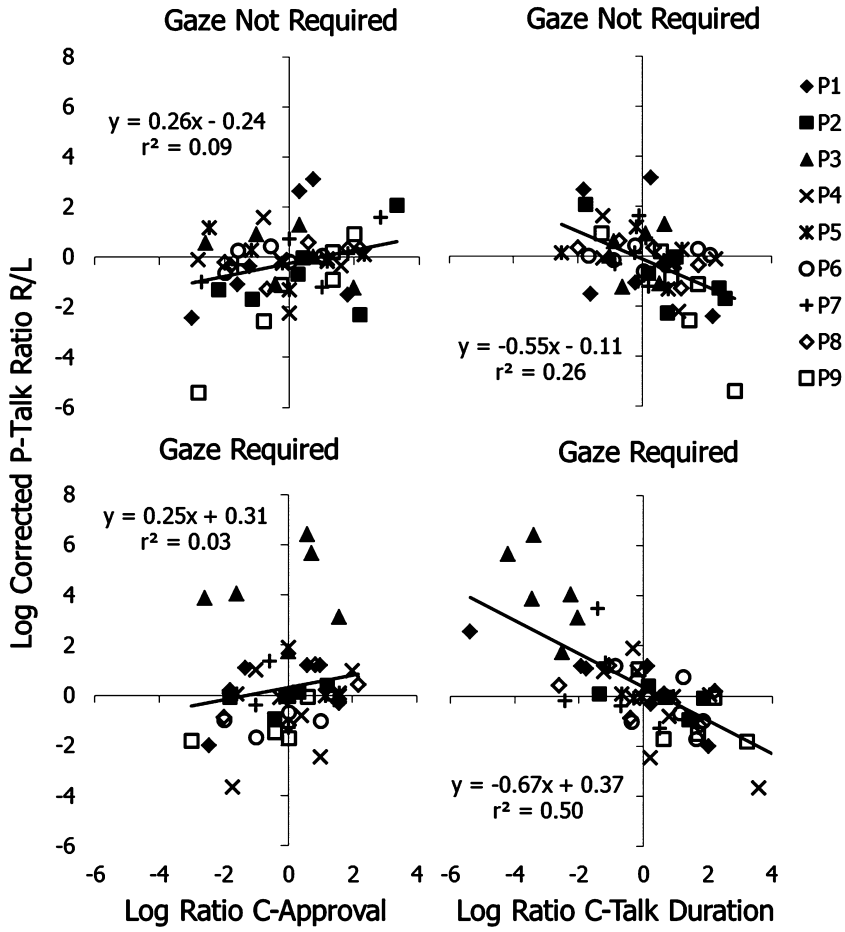


Fig. 3. Log ratio of participants' duration of talk to the left and right confederate corrected for bias, versus confederates' log ratio of approval (left panels) and confederates' log ratio of duration of talk (right panels). The two upper panels show corrected participants' talk ratios plotted against confederates' approval ratios (top left) or confederates' talk ratios (top right), when gaze was not required for approval. The two lower panels show the same relations when gaze was required for approval. The legend identifies the participants.

talk ratios. The two upper panels show corrected P-Talk ratios plotted against C-Approval ratios (left) or C-Talk ratios (right), when gaze was not required. The two lower panels show the same relations when gaze was required.

If relative approval were a good predictor for participants' relative talk, P-Talk ratios would follow a linear relation with a well-determined slope. Both conditions' r^2s shown in Figure 3 are less than .1, indicating that slopes were poorly determined. Low r^2s, together with points scattering around the x-axis, rather than a line of slope greater than zero, show that approval ratio exerted virtually no control over P-Talk ratio. Slopes in the upper and lower left panel are approximately equal, showing again that the contingency requiring gaze had no effect.

The two right-hand graphs show substantial negative slopes with higher r^2s, indicating that C-Talk ratio exerted substantial control over P-Talk ratio and that the relation was such that the participants spoke more to the confederate that spoke less, no matter if gaze was required.

Similar remarks apply to participants' relative distribution of gaze (P-Gaze ratio) shown in Figure 4. Approval ratio exerted little or no control over participants' gaze, whereas C-Talk ratio exerted substantial control over gaze in a way parallel to the control over P-Talk ratio;

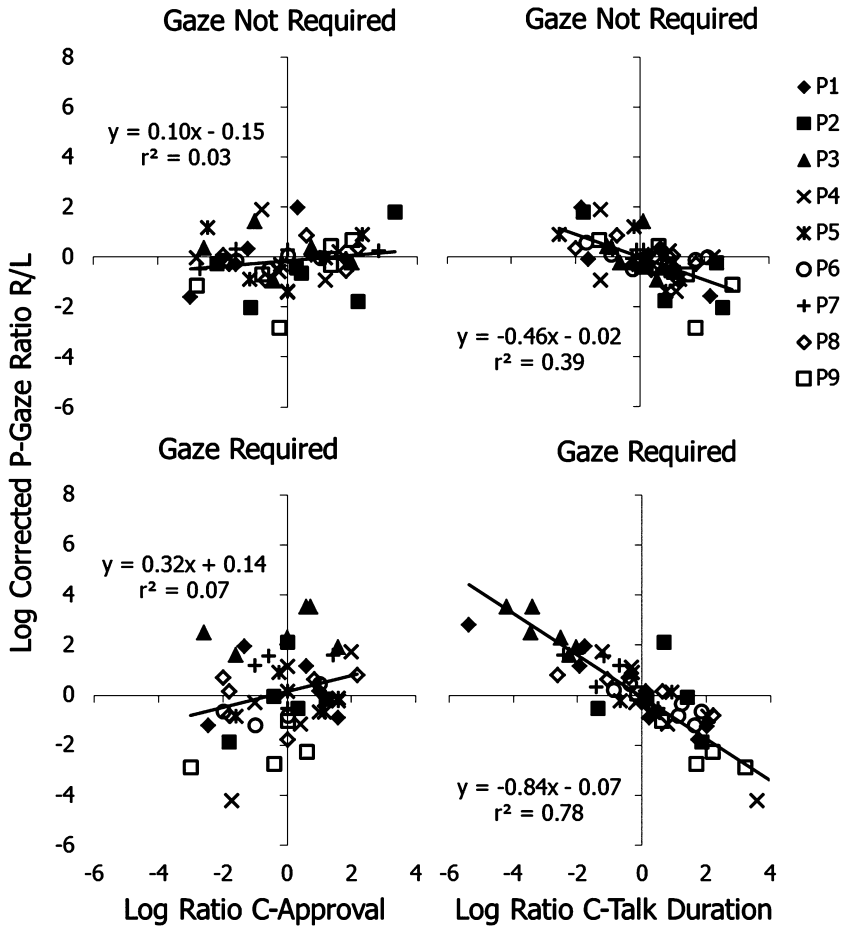


Fig. 4. Log ratio of participants' duration of gaze to the left and right confederate corrected for bias, versus confederates' log ratio of approval (left panels) and confederates' log ratio of duration of talk (right panels). The two upper panels show corrected participants' gaze ratios plotted against confederates' approval ratios (top left) or confederates' talk ratios (top right), when gaze was not required for approval. The two lower panels show the same relations when gaze was required for approval. The legend identifies the participants.

participants tended to look more at the confederate who spoke less.

Grouped Raw Data

To pool the raw data, counts and times were summed across participants, and then these sums were analyzed as if they came from one participant. Figures 5 and 6 show the results. Figure 5 shows the duration measures plotted against relative approval ratios. The top panel shows results for talk duration (TT_R/TT_L; "P-Talk"). A small positive slope of about 0.2 appeared for both conditions GN and GR. The lower panel shows results for gaze duration (TG_R/TG_L; "P-Gaze"). As with talk, a small positive slope appears, but smaller—about 0.1—and about the same whether gaze was required or not.

Figure 6 shows the same duration measures plotted against relative duration of confederates' talk. The vertical axes are the same as in Figure 5. Regression lines in Figure 6 confirm the negative relations of P-Talk ratio and P-Gaze ratio with C-Talk ratio. Three out of four of the slopes have r^2s greater than .8. Again, we see that participants favored in both speech and gaze the confederate who spoke less.

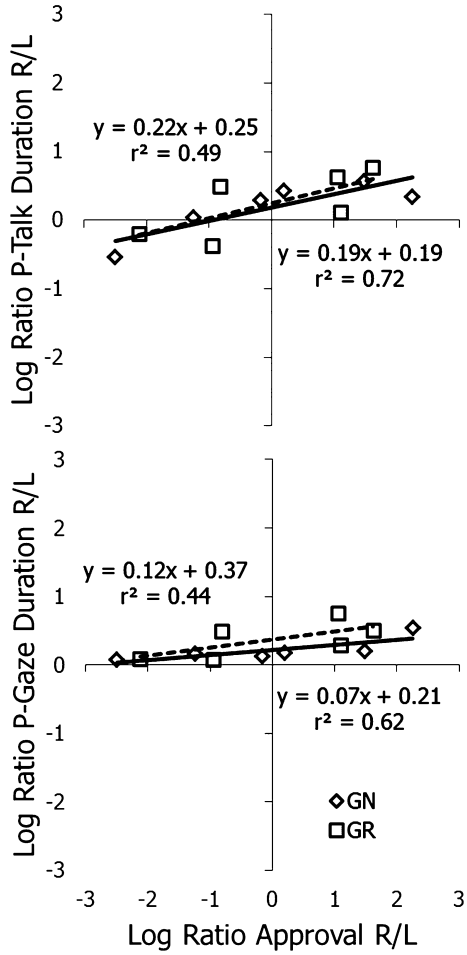


Fig. 5. Log ratio duration measures calculated from pooled raw data plotted against log approval ratios. The top panel shows results for relative participants' talk duration and the lower panel shows results for relative gaze duration. GN: gaze not required for approval. GR: gaze required for approval.

Linear Mixed Model Analysis for Repeated Measurements

We applied a linear mixed model (Everitt, 2009) to test the effect of the confederates' relative approval, their relative duration of talk and their relative frequency of talk on the participants' relative duration and frequency of talk and gaze. The linear mixed model is an extension to the generalized linear model, also referred to as multilevel model in the statistical literature. We used a linear mixed model with a random intercept for each participant to take the repeated measurements for each participant into account. Coefficients, *p*-values and confidence intervals are given in Table A3 in the Appendix. In line with our graphical analyses, relative approval did not significantly predict any of the dependent variables. However, the frequency with which confederates talked to the participant was a significant predictor for participants' behavior. All coefficients were negative, confirming the reverse relation between relative confederate talk and gaze and relative participant talk and gaze. Interaction analysis showed that conditions GR and GN do not significantly modify the predictors' effects, although overall smaller negative coefficients produced when eye contact was required suggest that confederates' behavior might generally have had a larger effect on participants' behavior when eye contact was required.

Changes in Matching During the Course of the Session

Borrero et al.'s (2007) data from the first 5 min of their sessions were better described by the matching relation than the data from the last 5 min. Conger and Killeen (1974) found the opposite. The degree to which the present participants' speech and gaze allocation conformed to Equations 2 and 3 did not differ substantially during different 5-min intervals of the sessions. All data presented here are based on performance during the last 5 min of each schedule, because participants were expected to need some experience with the current schedule of approval statements before it could take control over their behavior.

Whereas in an unpublished study (Milinder, 1997), in which one of the conditions resembled our GN condition, matching decreased as participants became more familiar during the course of the sessions, in the present study, the change of degree to which the participants matched their speech and gaze allocation to the confederates' approval statements did not follow a common pattern for all participants. Neither did the participants' topic preferences show any relation to the matching slopes they produced.

Participants' Overall Talk

Conger and Killeen (1974) excluded participants who did not speak in each minute of the

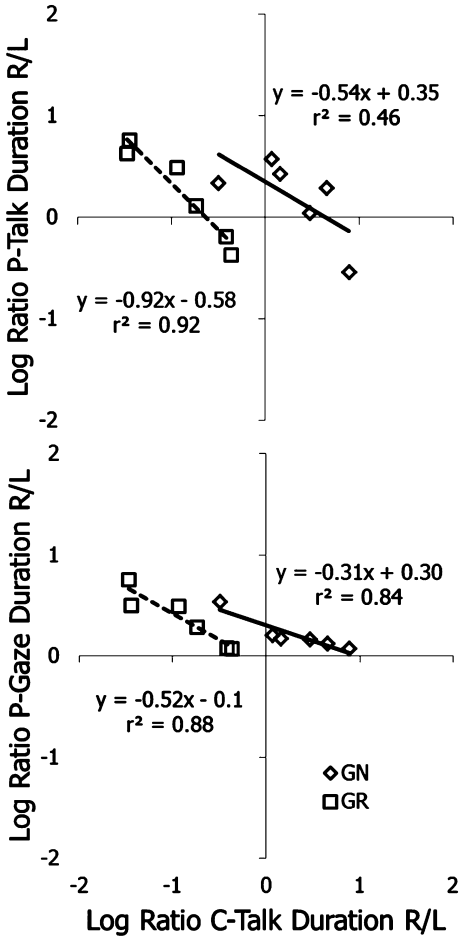


Fig. 6. Log ratio duration measures calculated from pooled raw data plotted against log duration of confederates' talk. The top panel shows results for relative participants' talk duration and the lower panel shows results for relative gaze duration. GN: gaze not required for approval. GR: gaze required for approval.

session from their analysis. None of the participants in the present study spoke in every consecutive minute, and only Participants 5 and 9 talked in every 2-min interval. Only Participants 6 and 8 looked at the confederate at least once in each minute. However, those participants who matched their speech and gaze allocation better to the confederates' relative rate of approval were not the ones that talked the most. There are no indications that our participants talked too little to display matching, as Figure 1 indicates. Female and male participants' behavior towards the confederates did not differ systematically in any of the analyses.

Discussion

No matter if individual data, pooled data, or raw data are analyzed, the results were the same: Little to no influence of relative approval on participants' talk and gaze (near-zero slopes in Figs. 2, 3, 4, & 5) and an inverse relation between confederates' and participants' relative talk (negative slopes in Figs. 2, 3, 4, & 6). These results contrast with those of Conger and Killeen (1974), which showed substantial influence of confederates' approval evidenced by a rough matching between relative talk and relative approval.

In the present experiment, we incorporated two conditions, one in which gaze was not required for confederates' delivery of approval (GN) and one in which gaze was required (GR; participant had to be looking at the confederate). Conger and Killeen (1974) did not require gaze for approval. We found no difference between the two conditions, and we cannot say what results Conger and Killeen would have obtained had they included the extra requirement.

Although our results differ from those of Conger and Killeen (1974), their finding of rough matching of relative talk to relative approval may indicate that their results were not due to strengthening by reinforcement, as seems often to be assumed, because gaze was not required to be contiguous with reinforcer delivery. A conventional notion of reinforcement would require that the behavior monitored for control (talk-plus-gaze) also be the behavior followed by reinforcers.

Both the earlier results apparently showing matching and our present results showing no difference whether gaze was required or not may suggest a different mechanism: what Segal (1972) and Baum (2012) have called *induction*. The simplest explanation of Conger and Killeen's (1974) results may be that their confederates' approval statements drew the participants' talk-plus-gaze in their direction—that is, talk-plus-gaze followed approval, not the other way around, as required by reinforcement. Thus, an approval statement induced a shift of talk-plus-gaze toward the confederate from which it came. In the present experiment, the disparity between the confederates' talk induced relatively less participant talk and gaze to the confederate who spoke more. An everyday explanation

would say that the participant was trying to draw the more taciturn confederate into the conversation, perhaps out of courtesy.

The confederates' comments differentially induced the participant's responding, but why did the allocation of talking and gaze remain largely insensitive to the difference in relative approval? Why was the participants' behavior sensitive to how much the confederates talked but not to the content (i.e., approval) of their statements? The reasons why some approval statements might not affect the participants' behavior can only be conjectured. The participants' history might have led to differential susceptibility to formally equivalent approval statements. The assumption that utterances, which we classify as approval, have an inducing or reinforcing effect on people's speech could not be tested in advance but could be assumed from previous research (Krasner, 1958). Of the approval statements delivered, some might have been ineffective due to differing establishing operations, defined as "any change in the environment which alters the effectiveness of some object or event as reinforcement" (Michael,1982; pp. 150–151).

As Kollins, Newland, and Critchfield (1997) pointed out after reanalyzing 25 studies employing variable-interval schedules of reinforcement, determining the role that establishing operations play in human matching behavior is central to an explanation of the variability among human studies as well as differences between human and nonhuman animals. Hence, one aim of our future research will be to gain a better understanding of the differential effect of establishing operations upon partners' relative utterances. Variables such as the current reluctance to build up new social contacts might differ among participants. Freshmen participants who have just moved to a new city to start studying might be more susceptible to others' approval or other statements regarding the topic under discussion than socially well-established or more mature participants would be. Incorporating such differences between participants or manipulating presumed establishing operations might prove valuable. Because establishing operations in the current study remain unknown and might vary across participants or across time, investigating possible systematic differences in setting or history that our participants were exposed to might be worthwhile. Perhaps, for example, the cultural background of our German participants arranged for a history in which integrating left-out conversational partners was reinforced.

Besides replication, the current study aimed to explore additional variables that can function as inducers in conversations. This exploration led to the observation that the confederates' relative rate and duration of overall speech covaried negatively with the participants' allocation of gaze and speech frequencies, as well as durations, to both conversational partners. To ensure the generalizability of this result to other participants and settings and to exclude the possibility that other unknown covarying variables are responsible for the results, we plan to test if we can systematically vary the duration and rate of the conversational partners' speech in future studies.

Inverse matching occurred even though the confederates were instructed not to differ in any behavior except their rate of approval. Confederates were instructed to talk as little as possible besides their approval delivery, and when one of them had made a comment, the other was to make a comment of equal length and content. Conger and Killeen's (1974) confederates "were limited to only cued interactions with the participant as described. They could make general statements directed to no one in particular, to each other, or to [the moderator]" (pp. 404). Restricting the confederates to the utterance of approval and striving to align their remaining behavior seem important because all of the confederates' activities, no matter how subtle, may affect the dependent variable. However, the effect of confounding variables needs to be weighed against the viability of the procedure.

Verplanck (1955) reported control over the utterance of participants' opinion statements in casual conversations using college student experimenters whose utterances were restricted to expressing agreement to those opinions. However, when Azrin, Holz, Ulrich, and Goldiamond (1961) attempted to replicate Verplanck's study, they discovered that most of their students who acted as experimenters in the replication were unable to follow the procedure and either falsified their data or had not followed the procedures properly but refrained from reporting their departure from instruction. To test if strictly

following the procedure was even possible, Azrin et al. trained four experimenters who held either a Master's or a Ph.D. degree in psychology, all well-trained in shaping of animal behavior, to engage in casual conversations with people while restricting their own comments to merely uttering agreement to opinions expressed by the person they talked to. The experimenters talked to 12 participants, all of whom left the room within 10 min, making it impossible for the experimenters to manipulate the participant's speech while restricting themselves to feedback on opinions.

Evidently, Conger and Killeen's (1974) confederates' behavior was not as restricted as Verplanck's (1955) and Azrin et al.'s (1961) but, although thoroughly trained, our confederates struggled to act as Conger and Killeen's confederates were supposed to do, that is, to balance their comments and to only talk to the participant when uttering approval. Because we do not know how well Conger and Killeen's confederates managed to follow the instruction of talking equally frequently and long, we cannot exclude the possibility that their confederates differed in talking rate and duration and that their utterances were negatively correlated to the participant's talk. All we know is that confederates' approval statements in Conger and Killeen's study showed a different relation to the participants' talk than they did in our study.

A possible difference between Conger and Killeen's (1974) confederates' and our confederates' ability to follow instructions could lie in the length of sessions. Our participants were exposed to two sessions lasting 180 min in total, and Conger and Killeen's participants conversed for a total of 30 min. Pierce et al. (1981) tested their six participants on 7 consecutive days for 1 hr per day. Their participants' behavior did not conform to a common pattern, but some produced results similar to our participants. Their participants' talking time and rate produced slopes ranging from $s = -.49$ to $s = .37$, displaying either undermatching, inverse matching or insensitivity to approval.

The effect of two conversational partners' agreement and overall speech upon a third person's speech is complex. The results of previous studies on matching in conversations range from inverse matching to severe undermatching and perfect matching to the absence of any pattern. Because those results render the applicability of the matching relation to conversations far from certain, forthcoming studies might start by examining the inducing effect (Baum, 2012) of a conversational partner's agreement and overall speech upon another conversational partner's speaking in a two-person setting. Such an approach would shed light on whether formally identical stimuli that have inducing properties in two-person conversations have the same function in three-person conversations.

As opposed to the content (in this case approval) of two conversational partners' speech, our results indicate that the relative duration of their utterances is a variable controlling a third conversational partner's speech. If this observation can be confirmed by systematic variation of the duration and frequency of confederates' speech as an independent variable, apart from approval statements, this finding could have wide-ranging implications for a number of social settings, such as therapeutic conversations, teamwork situations or everyday-life interactions between family members or colleagues. Besides creating the ground for those applications, our findings illuminate procedural possibilities to contribute to a scientific analysis of human verbal interactions that goes beyond Skinner's (1957) approach to the topic. Our test of applicability of the generalized matching relation (Eqs. 2 and 3) to conversational partners' allocation of speech and gaze represents one of many possibilities to enhance basic knowledge of conversational behavior by an analysis of speaking in the same general framework as other operant behavior.

References

Aparicio, C. F., & Baum, W. M. (2009). Dynamics of choice: Relative rate and amount affect local preference at three different time scales. *Journal of the Experimental Analysis of Behavior, 91*(3), 293–317.

Azrin, N. H., Holz, W., Ulrich, R., & Goldiamond, I. (1961). The control of the content of conversation through reinforcement. *Journal of the Experimental Analysis of Behavior, 4*(1), 25–30. doi:10.1901/jeab.1961.4-25

Baum, W. M. (1974). On two types of deviation from the matching law: Bias and undermatching. *Journal of the Experimental Analysis of Behavior, 22*(1), 231–242. doi:10.1901/jeab.1974.22-231

Baum, W. M. (1976). Time-based and count-based measurement of preference. *Journal of the Experimental*

Analysis of Behavior, 26(1), 27–35. doi:10.1901/jeab.1976.26-27

Baum, W. M. (1979). Matching, undermatching, and overmatching in studies of choice. *Journal of the Experimental Analysis of Behavior, 32*(2), 269–281. doi:10.1901/jeab.1979.32-269

Baum, W. M. (2012). Rethinking reinforcement: Allocation, induction, and contingency. *Journal of the Experimental Analysis of Behavior, 97*(1), 101–124. doi:10.1901/jeab.2012.97-101

Baum, W. M. (2013, November). *What counts as behavior.* Paper presented at the ABAI Conference on Theory and Philosophy, Santa Fe, New Mexico.

Baum, W. M., & Davison, M. (2014). Choice with frequently changing food rates and food ratios. *Journal of the Experimental Analysis of Behavior, 101*(2), 246–274.

Baum, W. M., & Rachlin, H. (1969). Choice as time allocation. *Journal of the Experimental Analysis of Behavior, 12*(6), 861–874. doi:10.1901/jeab.1969.12-861

Borrero, J. C., Crisolo, S. S., Tu, Q., Rieland, W. A., Ross, N. A., Francisco, M. T., & Yamamoto, K. Y. (2007). An application of the matching law to social dynamics. *Journal of Applied Behavior Analysis, 40*(4), 589–601.

Conger, R., & Killeen, P. (1974). Use of concurrent operants in small group research: a demonstration. *The Pacific Sociological Review, 17*(4), 399–416. doi:10.2307/1388548

Cuvo, A. J., Lerch, L. J., Leurquin, D. A., Gaffaney, T. J., & Poppen, R. L. (1998). Response allocation to concurrent fixed-ratio reinforcement schedules with work requirements by adults with mental retardation and typical preschool children. *Journal of Applied Behavior Analysis, 31*(1), 43–63.

Davison, M., & Baum, W. M. (2000). Choice in a variable environment: Every reinforcer counts. *Journal of the Experimental Analysis of Behavior, 74*(1), 1–24.

Everitt, B. S. (2009). *Multivariable modeling and multivariate analysis for the behavioral sciences.* CRC Press.

Fisher, W. W., & Mazur, J. E. (1997). Basic and applied research on choice responding. *Journal of Applied Behavior Analysis, 30*(3), 387–410. doi:10.1901/jaba.1997.30-387

Fleshler, M., & Hoffman, H. S. (1962). A progression for generating variable-interval schedules. *Journal of the Experimental Analysis of Behavior, 5*(4), 529–530. doi:10.1901/jeab.1962.5-529

Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior, 4*(3), 267–272. doi:10.1901/jeab.1961.4-267

Herrnstein, R. J., & Vaughan, W. (1980). Melioration and behavioral allocation. In J.E.R. Staddon (Ed.), *Limits to action: The allocation of individual behavior* (pp. 143–176). Academic Press.

Kollins, S. H., Newland, M. C., & Critchfield, T. S. (1997). Human sensitivity to reinforcement in operant choice: How much do consequences matter? *Psychonomic Bulletin and Review, 4*(2), 208–220. doi:10.3758/BF03209395

Krasner, L. (1958). Studies of the conditioning of verbal behavior. *Psychological Bulletin, 55*(3), 148–170. doi:10.1037/h0040492

Mazur, J. E. (2001). *Learning and behavior* (5th ed.) Upper Saddle River, NJ: Prentice-Hall.

McDowell, J. J., & Caron, M. L. (2010). Matching in an undisturbed natural human environment. *Journal of the Experimental Analysis of Behavior, 93*(3), 415–433. doi:10.1901/jeab.2010.93-415

Michael, J. (1982). Distinguishing between discriminative and motivational functions of stimuli. *Journal of the Experimental Analysis of Behavior, 37*(1), 149–155. doi:10.1901/jeab.1982.37-149

Milinder, L. A. (1997). *Matching in group conversation.* University of New Hampshire unpublished Master's thesis.

Pierce, W. D., Epling, W. F., & Greer, S. M. (1981). Human communication and the matching law. *Quantification of steady state operant behavior* (pp. 345–348). Amsterdam: Elsevier/North Holland.

Schlinger, H. D. (2008). The long good-bye: Why B.-F. Skinner's *Verbal Behavior* is alive and well on the 50th anniversary of its publication. *The Psychological Record, 58*(3), 329–337.

Schofield, G., & Davison, M. (1997). Nonstable concurrent choice in pigeons. *Journal of the Experimental Analysis of Behavior, 68*(2), 219–232.

Segal, E. F. (1972). Induction and the provenance of operants. In R. M. Gilbert & J. R. Millenson (Eds.), *Reinforcement: Behavioral Analyses* (pp. 1-34). New York: Academic.

Skinner, B. F. (1948). 'Superstition' in the pigeon. *Journal of Experimental Psychology, 38*(2), 168. doi:10.1037/h0055873

Skinner, B. F. (1957). *Verbal behavior.* Acton, MA: Copley Publishing Group.

Sloetjes, H., & Wittenburg, P. (2008, May). *Annotation by Category: ELAN and ISO DCR.* Paper presented at the 7th International Conference on Language Resources and Evaluation (LREC 2008), Marrakech, Morocco.

Thorndike, E. L. (1911/2000). *Animal intelligence: Experimental studies.* New York: The Macmillan Company.

Verplanck, W. S. (1955). The control of the content of conversation: reinforcement of statements of opinion. *The Journal of Abnormal and Social Psychology, 51*(3), 668.

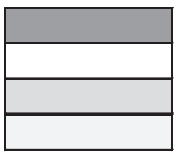
Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006, May). *Elan: a professional framework for multimodality research.* Paper presented at the Proceedings of 5th International Conference on Language Resources and Evaluation (LREC 2006), Genoa, Italy.

Received: May 12, 2016
Final Acceptance: February 8, 2017

Appendix

Table A1
All Individual Slopes Based on Performance During the Last Five Minutes of Each of the Six Schedules

Participant	Rate Approval Statements by Confederates							
	1 Duration talk		**2** Duration gaze		**3** Rate talk		**4** Rate gaze	
	s GN	s GR	s GN	s GR	s GN	s GR	s GN	s GR
1	0,65	0,44	0,36	0,00	0,30	–0,60	0,11	–0,12
2	0,40	0,20	0,30	0,61	0,23	–0,21	0,24	0,18
3	–0,46	0,13	–0,18	0,16	–0,06	–1,15	0,05	0,10
4	–0,15	0,53	–0,18	1,13	0,02	–0,83	0,10	0,66
5	–0,22	0,01	0,02	0,02	–0,04	0,02	–0,01	0,18
6	0,10	0,09	0,05	0,37	0,04	–0,38	0,07	0,18
7	0,31	–0,19	0,06	0,10	0,22	–1,02	0,12	–0,03
8	0,22	0,21	0,08	0,13	0,23	–0,03	0,08	0,06
9	1,22	0,32	0,40	0,29	1,01	–0,44	0,25	0,33
	Duration Talk by Confederates to Subject							
Participant	**5** Duration talk		**6** Duration gaze		**7** Rate talk		**8** Rate gaze	
	s GN	s GR	s GN	s GR	s GN	s GR	s GN	s GR
1	–0,61	–0,55	–0,59	–0,59	–0,42	–0,60	–0,18	–0,29
2	–0,76	–0,20	–0,65	–0,18	–0,49	–0,21	–0,41	–0,25
3	0,53	–1,39	–0,43	–0,86	0,12	–1,15	0,09	–0,74
4	–0,45	–1,07	–0,28	–1,23	–0,27	–0,83	–0,56	–0,75
5	–0,16	0,01	–0,50	–0,34	–0,13	0,02	–0,17	–0,29
6	0,06	–0,52	–0,11	–0,49	0,02	–0,38	–0,09	–0,16
7	0,02	–0,63	–0,22	–0,74	–0,23	–1,02	–0,15	–0,59
8	–0,33	–0,23	–0,30	–0,47	–0,38	–0,03	–0,26	–0,04
9	–1,44	–0,52	–0,62	–1,26	–1,09	–0,44	–0,48	–0,88
	Rate Talk by Confederates to Subject							
Participant	**9** Duration talk		**10** Duration gaze		**11** Rate talk		**12** Rate gaze	
	s GN	s GR	s GN	s GR	s GN	s GR	s GN	s GR
1	–0,82	–0,77	–0,79	–0,89	–0,56	–0,77	–0,24	–0,43
2	–0,80	–0,78	–0,71	–0,03	–0,54	–0,35	–0,44	–0,22
3	–0,02	–0,58	–0,80	–0,76	–0,14	–0,57	–0,02	–0,63
4	–0,28	–1,16	–0,06	–1,29	–0,25	–0,84	–0,50	–0,76
5	–0,41	0,04	–0,91	–0,50	–0,28	–0,04	–0,35	–0,55
6	0,08	–0,75	–0,15	–0,49	0,05	–0,31	–0,13	–0,14
7	–0,20	–0,43	–0,19	–0,66	–0,15	–0,90	–0,26	–0,58
8	–0,48	–0,63	–0,47	–0,73	–0,51	–0,49	–0,39	–0,50
9	–1,76	–0,51	–0,56	–1,71	–1,28	–1,01	–0,47	–1,23



$0.7 \leq s < 1.3$ indicating matching
$s < -1.3$ reverse overmatching
$-1.3 \leq s < -0.7$ indicating reverse matching
$-0.7 \leq s < 0.7$ indicating undermatching/no matching

Table A2
All Individual r^2 Based on Performance During the Last Five Minutes of Each of the Six Schedules

Participant	Rate Approval Statements by Confederates							
	1 Duration talk		**2** Duration gaze		**3** Rate talk		**4** Rate gaze	
	r^2 GN	r^2 GR	r^2 GN	r^2 GR	r^2 GN	r^2 GR	r^2 GN	r^2 GR
1	0,25	0,26	0,30	0,01	0,18	0,21	0,09	0,01
2	0,28	0,04	0,20	0,22	0,17	0,05	0,45	0,09
3	0,23	0,02	0,12	0,09	0,01	0,06	0,04	0,06
4	0,04	0,11	0,06	0,53	0,00	0,14	0,02	0,49
5	0,20	0,02	0,00	0,00	0,03	0,08	0,00	0,20
6	0,09	0,08	0,08	0,47	0,03	0,33	0,09	0,37
7	0,37	0,01	0,17	0,01	0,24	0,01	0,66	0,00
8	0,34	0,17	0,05	0,05	0,57	0,01	0,10	0,03
9	0,94	0,30	0,33	0,00	0,08	0,25	0,26	0,00
	Duration Talk by Confederates to Subject							
Participant	**5** Duration talk		**6** Duration gaze		**7** Rate talk		**8** Rate gaze	
	r^2 GN	r^2 GR	r^2 GN	r^2 GR	r^2 GN	r^2 GR	r^2 GN	r^2 GR
1	0,16	0,79	0,59	0,88	0,25	0,90	0,19	0,85
2	0,61	0,83	0,58	0,00	0,47	0,64	0,81	0,38
3	0,03	0,47	0,14	0,78	0,02	0,47	0,02	0,85
4	0,27	0,66	0,12	0,95	0,19	0,81	0,64	0,95
5	0,08	0,03	0,50	0,28	0,23	0,04	0,33	0,31
6	0,05	0,27	0,21	0,76	0,01	0,69	0,24	0,49
7	0,00	0,18	0,07	0,70	0,01	0,48	0,04	0,53
8	0,44	0,21	0,42	0,70	0,96	0,01	0,59	0,02
9	0,78	0,00	0,47	0,93	0,28	0,00	0,56	0,90
	Rate Talk by Confederates to Subject							
Participant	**9** Duration talk		**10** Duration gaze		**11** Rate talk		**12** Rate gaze	
	r^2 GN	r^2 GR	r^2 GN	r^2 GR	r^2 GN	r^2 GR	r^2 GN	r^2 GR
1	0,18	0,76	0,66	0,97	0,29	0,71	0,46	0,89
2	0,52	0,66	0,54	0,00	0,45	0,64	0,00	0,21
3	0,02	0,09	0,40	0,69	0,01	0,13	0,77	0,70
4	0,11	0,67	0,01	0,91	0,17	0,71	0,04	0,84
5	0,23	0,31	0,77	0,51	0,54	0,09	0,86	0,89
6	0,10	0,34	0,33	0,48	0,10	0,29	0,29	0,25
7	0,01	0,12	0,12	0,74	0,01	0,51	0,18	0,69

$r^2 \geq 0.7$
$r^2 < 0.7$

Table A3
Results of Multi-Level Regression Analysis

Participants' relative duration of talk	Coefficient	P-value	95% Conf. Interval
If gaze required	0.07	0.77	[–0.42, 0.56]
Confederates' relative duration of talk	–0.10	0.48	[–0.37, 0.17]
Confederates' relative frequency of talk	–0.59	0.00	[–0.94, –0.23]
Confederates' relative approval	–0.01	0.93	[–0.20, 0.17]
Participants' relative duration of gaze	Coefficient	P-value	95% Conf. Interval
If gaze required	0.10	0.51	[–0.20, 0.39]
Confederates' relative duration of talk	–0.29	0.00	[–0.45, –0.14]
Confederates' relative frequency of talk	–0.40	0.00	[–0.61, –0.19]
Confederates' relative approval	–0.10	0.09	[–0.22, 0.01]
Participants' relative frequency of talk	Coefficient	P-value	95% Conf. Interval
If gaze required	0.04	0.80	[–0.27, 0.35]
Confederates' relative duration of talk	–0.13	0.14	[–0.30, 0.04]
Confederates' relative frequency of talk	–0.44	0.00	[–0.67, –0.21]
Confederates' relative approval	–0.08	0.22	[–0.20, 0.05]