

Social Influences on Metacognition: Effects of Colearner Questioning on Comprehension Monitoring

Stuart A. Karabenick
Eastern Michigan University

Four experiments examined social influences on metacognition, testing whether learners' knowledge that coteachers have questions about material they are simultaneously viewing affects learners' own judged levels of comprehension. In Experiment 1 ($n = 88$), the frequency with which learners indicated they were confused increased with the number of questions they believed coteachers had about the material. Experiment 2 ($n = 38$) determined that the effect of coteacher questioning on self-judged comprehension was not due to distraction or social facilitation. Experiment 3 ($n = 100$) replicated the results of Experiment 1 and found that the social impact on learners' judgments of comprehension was less when questions were believed to have come from 3 coteachers rather than 1. Experiment 4 ($n = 60$) suggested that the number of questions per coteacher determines their impact on others' comprehension judgments.

Knowledge gains are more likely when learners are engaged, active, and self-regulating, using a variety of strategies, such as rehearsal, organization, elaboration, and critical thinking (Pintrich, 1989; Schunk & Zimmerman, 1994; Weinstein & Mayer, 1985). Effective learning also depends on metacognitive processes of planning, monitoring, and regulation that govern how and when strategies are best used (Paris, Lipson, & Wixon, 1983). Monitoring is essential to efficient knowledge acquisition, informing learners whether additional effort or alterations in strategies are required to accomplish planned activity. Whereas reliable and well-calibrated monitoring contributes to effective learning, inadequate or faulty monitoring places a limit on the effectiveness of strategy use (Flavell, 1979, 1987; Zimmerman, 1992, 1995; Zimmerman & Martinez-Ponz, 1992). For example, overestimating one's level of comprehension (overconfidence or the "illusion of knowing"; Bender, 1993; Epstein, Glenberg, & Bradley, 1984; Kulhavy & Stock, 1989) may result in the failure to engage in appropriate remediation (e.g., additional studying), leaving the learner underprepared. Underestimating comprehension can also be detrimental, resulting in unnecessary attention to what is already well understood.

Understanding the factors that affect comprehension

I thank John Knapp for his valuable comments on this article; James Todd, who developed the response interface; Donna Chapman and Priscilla Taylor, who helped develop and conduct the experimental sessions; Marsha Kolar for her usual editorial expertise; and Kumar Tadi for his assistance in data management. This research was supported in part by a Graduate School Research Fund Award and the Research on Teaching and Learning Program at Eastern Michigan University. Portions of Experiments 1 and 3 were presented at the annual meeting of the American Educational Research Association, Atlanta, Georgia, April 1993.

Correspondence concerning this article should be addressed to Stuart A. Karabenick, Department of Psychology, Eastern Michigan University, Ypsilanti, Michigan 48197. Electronic mail may be sent via Internet to psy_karabeni@online.emich.edu.

monitoring judgments would contribute to creating conditions that promote learning. Studies of the developmental, personal, and contextual determinants of monitoring have focused primarily on solitary individuals engaged in text comprehension (Baker, 1989; Jacobs & Paris, 1987; Markman, 1985; Paris et al., 1983) or studying (Zimmerman, Greenberg, & Weinstein, 1994). Because so much learning takes place in private, evidence from isolated individuals undoubtedly has considerable generalizability. However, because substantial learning also occurs in social settings (e.g., classrooms, small groups, and increasingly, computer-mediated environments), it is important to examine social influences on monitoring, adding to the expanding focus on social influences on cognition in general (e.g., Levine, Resnick, & Higgins, 1993). Although there are many similarities, metacognition in social contexts creates the potential for influences not present in solitary settings (Alexander, 1995). Examples of social influences on cognition are social comparisons (Festinger, 1954; Goethals & Darley, 1987; Suls & Wills, 1991) that affect learners' judgments of self-efficacy (Ruble, 1983; Ruble & Frey, 1991) and the way that co-actors can influence the interpretation of ambiguous stimulus events (e.g., Asch, 1955; Latané & Darley, 1970; Latané & Wolf, 1981; Sherif, 1935).

Social influences on comprehension judgments have been hypothesized but have yet to be demonstrated. Brown and Palincsar (1988), for example, proposed that cooperative learning arrangements can facilitate comprehension through their extension of "the locus of metacognitive activity by providing triggers for cognitive dissatisfaction outside the individual" (Glaser & Bassok, 1989, p. 644). In other words, one of the benefits of cooperative interactions is that learners may begin to wonder whether they understand material upon becoming aware that others have such doubts. Assuming that learners are typically under- rather than overprepared, socially induced uncertainty could be beneficial if it resulted in appropriate regulation. Although not considered by Brown and Palincsar, the obverse may also occur: The

failure of others to communicate their doubts could bolster one's own comprehension confidence. In this instance, learners would be making the following inference: "Everybody else seems to understand; perhaps I do too." This is analogous to the pluralistic ignorance that leads persons in groups to refrain from offering assistance in part because others' inaction influences their construal of events to suggest the absence of need (Latané & Darley, 1970; Miller & Prentice, 1994). Although there have been no studies of whether others' (here termed *colearners*) actions affect one's own monitoring judgments, the ubiquity of socially situated learning warrants examining whether such effects do, in fact, occur.

There are several ways that learners may signal their lack of comprehension (cognitive dissatisfaction; Darling, 1989), such as by direct statements to that effect (e.g., "Teacher, I don't understand") or indirectly by body language that requires an inference (e.g., a shrug or quizzical gaze). However, the most frequent way that students convey their ignorance or perplexity is by asking questions (Darling, 1989; Dillon, 1986). Colearner confusion could be attributed to the questioner's ability or motivation or to characteristics of the material itself (e.g., it is difficult or unorganized). In either case, because the presence of questions suggests to learners that others are confused, an increase is predicted in the likelihood that learners believe they are confused as well. Analogously, the absence of colearner questions should result in stronger inferences that colearners comprehend what is being presented, and increased confidence in their own level of comprehension. These hypothesized social influences on monitoring judgments were tested in three experiments that manipulated the frequency of colearner questions and the number of colearners believed to have asked them. A fourth experiment was designed to rule out whether effects of colearner questions on comprehension monitoring are attributable to factors unrelated to questioning.

Experiment 1

Although it would be desirable to test the hypothesized effects of colearner questions in contextualized settings, a number of factors preclude that approach, including (a) the difficulty of unobtrusively assessing comprehension monitoring judgments, (b) the problem that answers to colearners' questions could influence the degree of uncertainty that the questions's presence may have engendered, and (c) the complexity of isolating the effects of question asking from the myriad other determinants of monitoring in such contexts. Because the hypotheses focus specifically on the effects of learners' awareness of colearner questions, rather than their contents, even the process of asking questions could decrease confusion by providing critical information about the specific topic (e.g., "I know that electrostatic bonding requires ions, but does the octet rule always apply?") or increase it by communicating incorrect information (e.g., "Why does centrifugal force cause a rotating object to move toward the center of a circle?"). A controlled setting was therefore designed that provided learners with

information about the frequency with which colearners had questions about material they were simultaneously examining but conveyed no content.¹

Under the guise of a communications study, the degree of participants' failure to understand the contents of two videotaped messages (order counterbalanced within conditions) on a topic of contemporary interest (the environment) was assessed. All participants viewed the first message alone, and their responses during that interval were used as a baseline to control for individual differences in monitoring tendencies. Varying conditions were introduced during the second message presentation. Participants in a nonsocial control condition viewed the second message under the same individual conditions that existed during the first message. In three "social" conditions, participants were made aware of a colearner (simulated) who arrived in time to view the second message.² Social conditions differed in the number of questions signaled by the colearner during their simultaneous viewing of the second message.

Participants' judgments of their own comprehension inadequacy were predicted to vary in direct proportion to the frequency of colearner questions. The lowest level of confusion was expected when colearners never indicated having a question, and the most confusion was expected when colearners indicated having several questions. An intermediate condition was included to test whether a single question would be sufficient to affect monitoring judgments, representing the kind of impact that occurs, for example, in a classroom when one student breaks with a silent, nonquestioning majority (Morris & Miller, 1975). Inclusion of the nonsocial control condition made it possible to test whether the absence of questions decreases judged confusion or whether their presence increases it. Lower rates of confusion in the no-question social condition compared with the nonsocial control would indicate that the absence of questions leads learners to have greater confidence in their comprehension, whereas more confusion in the social questioning conditions than in the nonsocial control would indicate that questions decreased comprehension confidence.

Techniques used to assess comprehension monitoring have included nonverbal responses such as eye movements (Grabe, Antes, Thorson, & Hahn, 1987), reading time (Baker & Anderson, 1982), facial expressions and button pressing (Markman & Gorin, 1981), direct verbal statements, and ratings obtained either during or subsequent to task completion (Epstein et al., 1984; Garner, 1981). Evidence sug-

¹ Although this may seem unusual, becoming aware that others have questions without knowing their contents does occur frequently in classrooms, most notably when students signal their desire to ask a question (e.g., raise their hand) without being given the opportunity to complete their intention because they are not called on.

² The primary reason for using a simulated rather than an actual colearner was to control the frequency of signaled questions. In addition, the absence of direct contact with a colearner obviated the need to control for, or to balance, such colearner characteristics as gender, race, and age. Using real colearners would have added considerable variability, not to mention additional logistical complexity.

gests that these techniques are not equivalent (Markman, 1985). For example, verbal reports and rating scales that require retrospective judgments may not as accurately assess metacognitive activity that occurs during sustained periods of learning (Baker, 1989; Markman, 1985). Because a real-time indicator of monitoring was preferable to retrospective reports, the former was used to test for the hypothesized effects of colearner questions in the present study. The real-time indicator consisted of participant responses that indicated episodes of confusion. Retrospectively obtained verbal measures of comprehension during the presentation of materials were, nevertheless, included to determine if they would be consistent with real-time monitoring. It should be noted that an alternative real-time indicator response, question asking, was rejected because it would not be possible to distinguish the effects of colearner questions on monitoring from their influence on learners' question asking itself. In addition, question asking would have required inferring comprehension inadequacy.

I obtained additional retrospective ratings to examine participants' inferences about colearners from their question-asking behavior. Most important were inferences of colearner confusion. Because it is proposed that colearner questions imply colearner confusion, inferred colearner confusion should be directly related to the frequency with which they signaled having questions about the material. Other characteristics assessed were persistence, intelligence, motivation, and nervousness. We could be more confident that social influences on monitoring judgments were a function of inferred colearner confusion to the extent that these characteristics were unrelated to the frequency of colearner questioning. The degree to which participants were aware of their colearners was also assessed.

Method

Participants and Assignment to Conditions

Undergraduate psychology students (41 men and 47 women) volunteered and were paid \$5 to participate for 45 min in a communications study. In order of arrival, they were assigned to conditions using a block randomization procedure until there were 11 participants in each of eight groups that were formed from the four conditions and two order of message presentation subgroups in each condition. Within-gender assignment resulted in approximately the same ratio of men to women in each group. Preliminary statistical tests were conducted with gender of participant included. Because there were no main or interaction effects involving gender of participant in this or subsequent studies, gender effects are not reported. All sessions were conducted by the same female experimenter.

Stimulus Materials

Participants viewed two 8-min videotaped messages delivered by a male undergraduate student (a different student was used for each message). The contents consisted of information and opinions that were adapted from actual news articles about environmental issues. To ensure there would be an adequate level of confusion, the presentations were manipulated at 15 points by having the

presenter mispronounce words, omit sentences, speak too softly, rearrange words, or "garble" phrases. Pilot testing established that the alterations were not obvious and generated sufficient variability in how confused participants viewing the messages reported themselves to be. The order of presenting the messages was completely counterbalanced within conditions in all experiments reported here. Because the messages differed in difficulty, message order was included as a factor in all data analyses.

Physical Conditions and Procedure

Participants were seated individually in a room in front of a TV monitor and a partition that blocked visual access to the experiment's control and recording equipment. After preliminary introductions, the experimenter read the following instructions:

This study concerns communication. We are interested in the factors that influence the way that characteristics of the communicator and the communication are perceived by the viewer or listener. In order to investigate this process, we are asking students at Eastern Michigan University to provide us with their responses to videotaped material. We're interested in many aspects of these communications, and we are asking many people, such as yourself, to rate the communications along many different dimensions. Examples of these are: (a) emotional reactions to the communication, (b) how convincing is the communication, (c) whether and when a person has a question during the presentation, (d) whether and when a listener or viewer is confused or can't understand something during the presentation, or (e) the skill of the communicator. In your case we will be interested in the fourth on the list—when you feel confused or can't understand something.

Basically, we're going to present you with two videotaped presentations, both dealing with the environment, that last about 8 minutes each and ask you to indicate, by pressing a button, when you're confused or don't understand something. In order to make things more like a learning situation we're also going to ask you some questions about the material that was presented and about the presentation.

After participants read and signed an informed consent statement, they were given a small thumb-actuated button switch that had been calibrated to provide reliable response indications but required minimal pressure, after which the experimenter continued:

Pressing the button will indicate to us that you are confused or can't understand something that was said. Press the button and hold it down as long as you feel that way, then let it up. You may press the button as many or as few times as you wish during the presentation to indicate when you feel you are confused. Please hold down the button for as long as you feel confused, then release it.³

To relate this to something you may do frequently, just pretend that you are in a classroom and the teacher is teaching or explaining something to you. Sometimes you understand

³ The amount of time that participants held their button down was recorded but not analyzed because this measure was apparently less reliable than the frequency of responses, and it was not sufficiently sensitive to condition effects to be statistically detected. It should be noted that within-group correlations between frequency and total time [$\ln(\text{time in seconds} + 1)$] were moderately positive, with values ranging from the mid .30s to the high .60s.

what's being said, other times you do not. We're not interested in how you think you should react, but rather how you might actually respond in a setting such as the classroom.

Before we begin, let's practice on some sentences that I'll read to you. Some of them will be easier to understand than others. Just press the button when you don't understand what's being said in the sentences as I read them.

The experimenter then read a list of 10 sentences designed to elicit confusion and recorded the number of button presses. After ensuring that the participant understood the instructions and how to indicate confusion, she continued:

O.K., now we'll show you the presentation that will be displayed on this monitor. In order to standardize the situation and remove any distractions, I'm going to turn off the overhead lights. There will be a blank space on the tape and then the presentation will begin.

Baseline assessment. The first message was then displayed and the onset and duration of each button press was automatically recorded. Each press also actuated a light on the experimenter's console behind the partition, its glow being visible to the participant as a result of the low ambient illumination. The light was designed to be noticeable but unobtrusive and did not interfere with the ability to see the material being presented on the monitor. The participant's light signal was used to add credibility to the light signal presumably activated by the colearner during the second message presentation. At the completion of the first video presentation, I tested participants on its contents, using a series of multiple-choice questions, to help maintain the cover story.

Conditions. In the nonsocial control condition, the experimenter retrieved the test and began the second presentation after a reminder that "Your response is just as before. Press the button when you are confused or can't understand something."

In the social conditions, participants were led to believe there was another student (colearner) who simultaneously viewed the second message. All social conditions were identical except for the number of questions signaled by the colearner who was presumably in an adjacent room. After presenting the first multiple-choice quiz, the experimenter stated that she "wanted to see if another subject that was expected had arrived" whom she would "get started" before returning to present the second video.

The experimenter then left the room, closed the door to the hall, and role-played first greeting and then ushering the "other student" into the adjacent room. She remained there and presented to that student the same instructions read to the participant, including a request that the student sign the informed consent statement. Physical conditions were such that participants could hear, albeit faintly, the experimenter's instructions to the other student. Upon returning to the participant's room the experimenter collected the quiz and continued:

Just to let you know what's going on, there is another subject in the other room who will be watching the second video along with you. That person is going to indicate to me when there is a question about the material. That person is the 'question' condition. You are still in the 'confusion' condition. When you feel confused, you are to press the button.

Thus, the instructions reminded participants that they were still in the confusion condition and that they were about to view the second, and last, message.

To further ensure that the participant understood the conditions they and their colearners were in, the experimenter placed small

signs on a condition indicator board that participants could easily see. To complete the dissimulation, the experimenter presumably communicated with the other participant through a microphone, asking that his or her button be pressed to indicate readiness to proceed. The simulated response was indicated by a tone and the onset of a green light on the experimenter's console (the participant's was red), the glow of which was also visible above the separating partition.

Colearner questioning frequency. The incidence of colearner responses that signaled they had questions to ask varied by condition. In one condition (S0), colearners never indicated they had a question during the entire presentation. In another condition (S1), a single response occurred 15 s into the presentation. Colearners signaled having 12 questions in a third social condition (S12). To increase the dissimulation's credibility, colearner questions were programmed to follow, with varying delays, the deliberate distortions introduced into the messages. For example, colearner questions were signaled after the following number of seconds into one of the message presentations: 15, 25, 53, 82, 132, 175, 215, 280, 300, 330, 382, and 415.

Following the second presentation, participants were once again given a brief quiz on its contents to remain consistent with the cover story, whereupon the experimenter entered the adjacent room to simulate the same procedure for the colearner. On completion of the quiz, participants, and presumably, colearners were given a postexperimental questionnaire that provided manipulation checks and other retrospective ratings.

Postperformance retrospective ratings. Portions of the postexperimental questionnaire asked participants whether and when there was another student viewing the presentation in another room. All those in the nonsocial control condition and in each of the social conditions responded appropriately. After indicating there was another student, those in the social conditions were then asked to indicate how "aware" they were of the other student, responding on a 10-point (0 to 9) scale with anchors of *not at all* and *very much*. To verify that participants in the social conditions (S0, S1, and S12) understood the task of their simulated colearner, we asked the participants to select their colearner's task from among the five conditions that were listed in the initial instructions.

All participants in the social conditions selected the appropriate alternative ("That they had a question during the presentation"). All those in the social conditions also appropriately selected the "does not apply" option when asked, "About how many times during the *first* presentation did the person in the next room press their button to communicate to the experimenter?" When asked about button presses during the second presentation, 20 of 22 in the S0 condition reported appropriately that the coparticipant never asked a question (the others reported one question but were retained in the analyses). Most of those in condition S1 overestimated the number of colearner questions asked, with modal categories of two and three. Participants in S12 tended to underestimate the other's questions, with most participants (17 of 22) estimating between 7 and 10. Despite the less than perfect encoding and recall of colearner questions, the distributions had virtually no overlap. The means of 0.09 ($SD = 0.3$), 2.5 ($SD = 2.4$), and 8.5 ($SD = 1.0$) in S0, S1, and S12, respectively, varied statistically, $F(2, 60) = 179.00, p < .0001, MSE = 2.31$, primarily attributable to the linear trend, $F(1, 60) = 342.73, p < .0001$. All means were statistically different from each other (all $ps < .01$).

Participants in all conditions also provided retrospective judgments of how confused they were during the first and second messages using three response formats. The first, subsequently referred to as the *Likert scale* measure, consisted of four statements

and a 6-point (0 to 5) response format with anchors of *not at all like me* and *very much like me*. The statements were "I was not confused at all during the first (second) presentation," "Most of the time I didn't understand what the first (second) person was saying," "The first (second) presentation seemed clear to me," and "There were many times that I couldn't follow the first (second) presentation." Mean responses over the four items were computed after reversing ratings on the second and fourth item. The second format, which is referred to subsequently as the *self-rating*, asked participants to indicate directly their degree of confusion by placing an F (to represent the first presentation) and S (to represent the second presentation) on a single 10-point (0 to 9) dimension anchored by *was very clear to me* and *was very confusing to me*. The third format, referred to as the *percentage of time confused* asked for the percentage of time that participants felt they were confused during each of the presentations.

Within-group correlations were computed between the real-time frequency of confusion responses during each message presentation and the respective postexperimental ratings of confusion. The correlations were then averaged over the eight groups using *r* to *Z* transformations and are shown in Table 1. The self-ratings using all rating formats correlated significantly ($p < .001$, with 64 *dfs*) with their parallel real-time response frequencies: the number of responses during the first presentation and ratings of confusion during the first presentation, and the number of confusion responses during the second presentation and ratings of confusion during the second. Furthermore, correlations were relatively low between nonparallel ratings and response frequencies: the number of responses during the first presentation and ratings of confusion during the second presentation, and the number of confusion responses during the second presentation and ratings of confusion during the first. Thus participants' retrospective judgments of their experiences when viewing the presentations are consistent with their responses during the presentations.

Results and Discussion

Real-Time Monitoring

The mean frequency of confusion responses for all conditions combined was 4.1 ($SD = 4.0$) during the first message presentation and 3.7 ($SD = 3.8$) during the second. Mean confusion responses during the baseline were 5.3

($SD = 4.1$), 4.3 ($SD = 3.7$), 3.8 ($SD = 4.8$), and 3.2 ($SD = 3.4$) for the nonsocial control condition and S0, S1, and S12 conditions, respectively. (Note that, although it is not necessary to test for equivalency of the baseline means because of random assignment, they did not vary statistically: $F[3, 83] = 1.03$, $p > .05$, $MSE = 15.07$). The mean frequencies of responses for those conditions during the second condition were 4.3 ($SD = 4.5$), 2.8 ($SD = 2.9$), 3.3 ($SD = 3.2$), and 4.5 ($SD = 4.4$). Change from baseline (frequency during second presentation minus frequency during the first) was used to test for effects of colearner questioning frequency.

Figure 1 presents the mean changes in the frequency of confusion responses from the first to the second message presentation for the three social conditions, with the nonsocial control as a point of reference. As hypothesized, the number of judged episodes of confusion during the second presentation (controlling for the baseline frequency) was a monotonic function of the number of questions the colearner was presumed to have asked. Furthermore, the nonsocial control condition mean of -1.0 was between that of conditions S0 and S1, suggesting that the presence or absence of questions affects comprehension judgments compared with conditions where social comparison information is not available. Statistical analysis of change scores was conducted in two stages. The first, which included only the three social conditions, used a 3 (frequency of colearner questions) \times 2 (order of message presentation) analysis of variance (ANOVA) to test the relationship between confusion responses and the frequency of colearner questions. The linear trend was used to test the predicted monotonic effect of questioning on monitoring judgments (see Rosenthal & Rosnow, 1985; Rosnow & Rosenthal, 1988). Using weightings appropriate for the unequally spaced frequency of question intervals (-0.46 , -0.35 , and 0.81 for S0, S1, and S12, respectively), the linear trend was statistically significant, $F(1, 60) = 6.60$, $p < .02$, $MSE = 11.81$, which supports the hypothesized influence of colearner questioning.

A 4 (frequency of colearner questions) \times 2 (order of message presentation) ANOVA and directional Dunnett tests ($\alpha = .05$) were used to determine whether colearner questions affected participants' judgments of confusion in the predicted directions when compared with the nonsocial control. The frequency of colearner questions effect was statistically significant, $F(3, 80) = 2.76$, $p < .05$, $MSE = 11.38$. In addition, the mean of condition S12 was statistically greater than that of the nonsocial control condition, indicating that colearners who asked several questions increased participants' doubts about their own comprehension of the material being presented. However, although in the predicted direction, the mean for condition S1 was not statistically higher than in the control, suggesting that a single question was insufficient to increase participant confusion. Nor was the mean of condition S0 statistically lower than that of the nonsocial control; thus there is no statistical justification in the present experiment for concluding that a colearner who asks no questions lowers self-judged confusion, compared with a condition with no social comparison

Table 1
Correlations Between Real-Time Confusion Responses and Retrospectively Reported Confusion (Experiment 1)

Rating format	Responses during message presentation	
	First	Second
First presentation		
Likert scale	.52**	.10
Self-rating	.65**	.05
% of time confused	.53**	.10
Second presentation		
Likert scale	.01	.62**
Self-rating	-.03	.59**
% of time confused	.10	.61**

Note. Values shown are averages, using *r* to *Z* transformations, of correlations within condition and message order subgroups.

** $p < .001$. ($df = 64$).

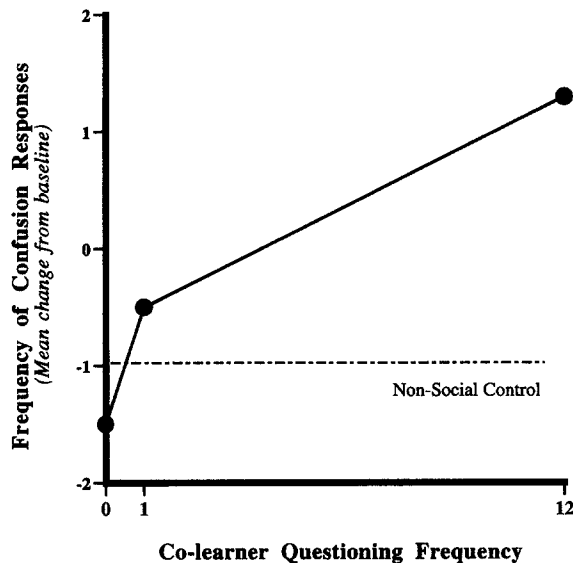


Figure 1. Mean number of participant confusion responses (change from baseline to testing) as a function of colearner questioning frequency conditions in Experiment 1.

information. It should be noted that the amount of change from baseline to testing in condition S0 was only one-third of the distance from the average of 4.1 confusion responses during baseline to a possible level of 0 during the second message presentation. Therefore, it is unlikely that the statistically nonsignificant difference between conditions S0 and NS was due to a floor effect.

Mediating Process Analysis

Mean postexperimental ratings, shown in Table 2, were used to examine the relationship between participants' inferences of colearner characteristics and the frequency of questioning. A series of 3 (frequency of colearner questions) \times 2 (order of message presentation) ANOVAs on retrospective ratings in the social conditions indicated that only inferred colearner confusion and persistence were influenced by colearner questioning frequency. Colearner confusion varied significantly among conditions, $F(2, 60) =$

25.13, $p < .0001$. On the basis of Fisher multiple-comparison logic (Levin, Serlin, & Seaman, 1994), a linear trend test was conducted with appropriately weighted contrast coefficients. As predicted, that test was statistically significant, $F(1, 60) = 41.25$, $p < .0001$, $MSE = 3.94$. In addition, all three means were significantly different from each other (all $ps < .01$). Thus, just one colearner question was sufficient to increase participants' inferred level of colearner confusion (compared with colearners who asked no questions).

Additional evidence that inferred colearner confusion affected judged comprehension comes from relationships between ratings that assessed these two dimensions. Averaged within-group correlations (using r to Z transformations, $df = 48$) between inferred colearner confusion and participants' self-perceived confusion during the second message presentation tended to be correlated: Likert format, $r = .27$, $p < .05$; self-rating, $r = .24$, $p < .05$; percentage of time confused, $r = .20$, $p < .10$. All of the correlations were significant (at $p < .05$) after partialing out all other inferred dimensions. Thus, the more that participants perceived their colearners as confused, the more confused they rated themselves.

Mean perceived colearner persistence also differed significantly among conditions, $F(2, 60) = 10.28$, $p < .0002$, $MSE = 5.54$, which is attributable primarily to the difference between condition S12 and each of the other two conditions, S0 and S1 (Newman-Keuls, $p < .01$). This suggests an alternative manner in which the frequency of colearner questioning may have affected participants' judged comprehension inadequacy. If this were the case we would expect, analogous to the analysis above, an association between perceived colearner persistence and self-reported confusion. However, an examination of the averaged within-group correlations between perceived persistence and participants' confusion responses does not support this alternative. All averaged correlations were low and nonsignificant: Likert format, $r = .02$; self-rating, $r = .08$; percentage of time confused, $r = .01$. Thus, inferred levels of colearner confusion, but not other inferred characteristics, appear to be linked to participants' own monitoring judgments.

Participants' Awareness of Colearners

An additional factor that could have contributed to the differences between frequency of colearner questioning conditions is participants' awareness of colearners. For example, the difference between S1 and S12 could be attributed to participants' having been less aware of a colearner who asked one question than a colearner who asked 12 questions. Furthermore, the small (and statistically nonsignificant) difference in real-time confusion responses between conditions S0 and S1 could have occurred because participants were not sufficiently aware of the colearner in condition S1. A 3 (frequency of questioning) \times 2 (order of message presentation) ANOVA of awareness ratings yielded a significant frequency main effect, $F(2, 60) =$

Table 2
Mean Inferred Characteristics of Colearners in Social Conditions (Experiment 1)

Dimension	Colearner question frequency			$F(2, 60)$	p
	0	1	12		
Confused	2.1 _a	4.2 _b	6.5 _c	25.13	.0001
Persistent	3.6 _a	4.0 _a	6.7 _b	10.28	.002
Intelligent	5.9	6.2	5.2	1.54	ns
Motivated	5.0	5.2	6.2	1.66	ns
Nervous	4.0	4.8	5.0	<1.00	ns

Note. Ratings are based on a 0 to 9 scale. Noncommon letter subscripts within rows indicate means that are significantly different by post hoc pairwise comparisons (based on $\alpha = .01$).

15.09, $p < .0001$. The linear trend associated with frequency was also statistically significant, $F(1, 60) = 11.11$, $p < .001$, $MSE = 5.78$, and additional comparisons between means (at $\alpha = .01$) indicated that participants in S0 were not as aware of their coparticipant ($M = 1.3$, $SD = 1.9$) as were participants in condition S1 ($M = 4.7$, $SD = 2.5$) or S12 ($M = 4.9$, $SD = 2.8$), which did not differ from each other. It follows that the minimal difference in self-judged confusion between S0 and S1 was apparently not due to participants being unaware of their coparticipant in condition S1, and the difference in judged colearner confusion between conditions S1 and S12 is not attributable to differences in colearner awareness.

Retrospective Self-Reported Confusion

Also examined was whether the frequency of colearner questions would affect retrospectively reported confusion as it did real-time confusion responses. A series of linear trend tests from 3 (frequency of colearner questions) \times 2 (order of message presentation) ANOVAs on Likert format, self ratings, and percentage of time confused (changes from baseline to testing) were all nonsignificant ($F_s < 1$). Thus, participants' retrospective reports did not reflect the frequency of colearner questioning differences that were found using real-time confusion responses.

In summary, evidence supports the hypothesized relationship between another's questioning behavior and one's own judgments of comprehension adequacy. More frequent colearner questions appear to increase the incidence of judged comprehension inadequacy during message presentations. Comparisons with a control condition that provided no social comparison information support the conclusion that more frequent questioning increases doubts about one's own comprehension, but not its inverse—that the absence of questions increases comprehension confidence. In other words, there is evidence that being aware of others who ask questions can trigger cognitive dissatisfaction (Glaser & Bassok, 1989) but not that the absence of questions increases cognitive satisfaction. There is also evidence indicating that the effects of colearner question-asking on monitoring judgments is mediated by inferred levels of colearner confusion.

Experiment 2

Before concluding that colearner questioning and the inferred level of colearner confusion affect self-monitoring judgments, I conducted a second experiment to rule out two plausible rival hypotheses. One is that colearner questions affected participants' own judged comprehension because they interfered with the actual comprehension of the messages. Despite attempts to minimize distraction and ensure the ability of participants to detect the presence of others' questions, the stimuli used to signal colearner questions could have interfered with task performance. Thus, participants may have indicated they were more confused when others signaled having more questions because they were

distracted, finding it more difficult to understand the material being presented to them, not because of inferences about their colearner's level of comprehension.

A second possibility derives from social facilitation theory (Bond & Titus, 1983; Zajonc, 1965). Applied to the present context, although evaluation apprehension is unlikely (colearners were presumably not aware of participants' performance; Cottrell, Eisenberger, Sekerak, & Rittle, 1968), the mere (implied) presence of colearners could have increased arousal. Higher levels of arousal would, in turn, be expected to increase the likelihood of the simple indicator response used to signal confusion. Thus, a higher rate of confusion responses would be predicted for conditions in which participants were aware of colearners, compared with conditions where they were relatively unaware. Because the results of Experiment 1 indicated that the awareness of colearners was higher in condition S12 than in S0, the differences in confusion could be attributed to differences in arousal rather than, as was concluded, inferred colearner confusion.

To test whether distraction or social facilitation could have produced the effects of colearner questioning frequency on judged confusion, the stimuli used to signal colearner confusion were presented exactly as they were in conditions S0 and S12 in Experiment 1, but were described to participants as indicative of a colearner state other than having a question about the material. If the results found in Experiment 1 were replicated without the link between the stimuli and colearner questions, then the effects found in Experiment 1 could have been a function of distraction or social facilitation. Alternatively, the failure to replicate the results of Experiment 1 would suggest that neither distraction nor social facilitation were operative and that differences in signaled confusion as a function of colearner questioning frequency are attributable to differences in participants' inferences of the degree of colearner confusion.

Method

Participants and Procedure

Participants (17 men and 19 women) were from the same population used in Experiment 1. They were randomly assigned to four groups (conditions S0 or S12 and two message orders within conditions) in the same manner as in the previous experiment. All procedural elements of the study were likewise identical, with the exception that participants were informed that colearner responses indicated that they had "an emotional response to the material being presented" rather than a question. This established conditions that would create differences in levels of awareness comparable with those when colearner responses signified having questions. Because having an emotional response was in the set of colearner conditions described to participants in Experiment 1, manipulation checks and postexperimental assessment could remain unaltered.

Manipulation Check

Comparability with Experiment 1 is suggested by participants' estimates of the frequency of colearner emotional responses. Spe-

cifically, the mean estimated frequency of colearner responses was 0.05 ($SD = 0.2$) in condition S0 and 8.6 ($SD = 2.9$) in condition S12, $F(1, 32) = 143.17$, $p < .0001$, $MSE = 1.36$, whereas the values in Experiment 1 were 0.09 and 8.7, respectively.

Results and Discussion

A 2 (frequency of colearner emotional reactions) \times 2 (order of message presentation) ANOVA of participant confusion responses (change from baseline) yielded no evidence that the number of signaled coparticipant emotional reactions affected participants' judgments of confusion. Not only was the frequency main effect negligible ($F < 1$) but the mean changes from baseline were virtually identical at $-.33$ ($SD = 2.6$) in condition S0 and $-.39$ ($SD = 4.0$) in condition S12. In addition, the absence of comprehension monitoring effects occurred despite a difference in participants' awareness of their colearnings as a function of frequency of emotional responses, $F(1, 32) = 78.56$, $p < .0001$, $MSE = 3.40$, which was comparable with that obtained in Experiment 1. Mean awareness ratings were 0.3 ($SD = 0.8$) in condition S0 and 5.7 ($SD = 2.4$) in condition S12 (compared with 1.3 vs. 5.7 in Experiment 1). Analyses of other postexperimental ratings indicated that participants perceived colearnings who signaled 12, compared with those who never signaled having an emotional response, as more motivated ($M = 6.1$, $SD = 2.6$ vs. $M = 3.8$, $SD = 1.6$), $F(1, 32) = 10.52$, $p < .01$, $MSE = 4.67$, and persistent ($M = 5.7$, $SD = 2.6$ vs. $M = 3.4$, $SD = 2.3$), $F(1, 32) = 7.70$, $p < .01$, $MSE = 6.40$. However, colearnings were considered neither more nor less intelligent, nervous, or confused.

The evidence thus suggests that the influence of colearner questioning frequency on participant confusion found in Experiment 1 was not a function of either distraction or social facilitation. This implies that participants' own monitoring judgments were due to colearner questioning frequency and participants' inferred level of colearner confusion. With increased confidence that the effect is not artifactual, additional experiments were conducted to determine whether it would replicate with an independent sample and the moderating effect of questioning by multiple colearnings.

Experiment 3

Because social learning contexts such as classrooms and many peer learning opportunities are typically more than dyadic, we next examined whether effects of questions on monitoring judgments would be affected by the presence of additional colearnings. That additional colearnings signaling they are confused would be expected to have a more substantial effect on judged confusion seems intuitive, analogous to the difference between one student raising his or her hand to ask a question and several students doing so. We would also expect that result on the basis of social impact theory (Jackson, 1987; Latané, 1981), which proposes that the impact of such processes as persuasion is greater as the number of persuaders increases. From an attribution perspective (Kelley, 1967; Weiner, 1979, 1985), more colearn-

ers can be viewed as providing additional consensus information. That is, the frequency of colearner questioning provides stronger evidence that the material (i.e., the entity) learners are attempting to understand is difficult. As noted earlier, passive bystanders reduce the likelihood of intervention in part because they influence beliefs and provide consensus information about the nature of the situation (i.e., that it is not an emergency), an effect that is stronger with several than with one other passive respondent (Latané & Darley, 1970). In the present instance, several nonquestioners who presumably understand the material should be more influential in convincing others that the material being presented is comprehensible. Thus, the effect on self-judgments is predicted to be greater when nobody in a larger group has questions.

To test the hypothesis that the impact of questions would be greater with more colearnings, two additional conditions were added to conditions S0 and S12 from Experiment 1: one in which 3 colearnings either had no questions (condition 3S0), and another in which 3 colearnings indicated a total of 12 questions between them (condition 3S12). This produced a 2 (frequency of colearner questions) \times 2 (number of colearnings) \times 2 (order of message presentation) factorial design. A nonsocial control condition was included that, in conjunction with the single colearner conditions (S0 and S12), permitted testing whether the effects found in Experiment 1 would replicate.

Method

Participants, Assignment to Conditions, and Procedure

Participants in this experiment (40 men and 60 women) were from the same population and volunteered under the same general conditions as those in Experiments 1 and 2. In order of arrival, they were placed by the experimenters (two were involved in this experiment) into conditions using a block randomization procedure until they had each assigned 5 participants to each of the 10 groups formed by the five conditions and two message order subgroups. To control for possible differences, I included an experimenter dimension in all data analyses. Once again, within-gender assignment was used to produce approximately the same ratio of men to women in each condition. The stimulus materials and the general procedure were identical to those in Experiments 1 and 2 with one colearner (S0 and S12), differing only as described below for participants in conditions with 3 colearnings (3S0 and 3S12).

Procedure in Multiple Colearning Conditions

In the conditions with 3 colearnings, after participants viewed the first message and completed the quiz on its contents, the experimenter left the room and presumably greeted the arriving students by asking rhetorically if "All three of you are here for the communications study?" She then simulated ushering them into the adjacent room, stating

Since all three of you will be under the same conditions, I will explain to you in here what you will be doing; then one of you

will stay here and I'll take two of you down the hall into different rooms.

She read the instructions, which could be overheard by the participant, and then role-played directing 2 of the colearners to other rooms. After returning, she explained to the participant that the 3 "subjects" in the other rooms were all in the questioning condition, and would be pressing their buttons to indicate that they had a question to ask. This was reinforced by placing condition identification signs on the conditions display board that participants could clearly see. She then communicated to each of the 3 by microphone, asking them, in turn, to actuate their buttons, which produced easily discriminable tones (although the same light signal). The remainder of the procedure was identical to Experiment 1, with modifications required by having 3 colearners rather than 1, such as minor alterations on the postexperimental questionnaire.

Estimated Colearner Questioning Frequency

As in Experiments 1 and 2, all participants appropriately identified the correct colearner conditions in place during presentations of the first and second messages. Also similar to the previous experiments, participants' estimates of the frequency of colearner questions were clearly differentiated: frequency of questions main effect, $F(1, 64) = 147.87$, $p < .0001$, $MSE = 6.37$, with mean estimates of .1 ($SD = .3$) in the no-question conditions (S0 and 3S0 combined), and 9.8 ($SD = 4.6$) in the 12-question conditions (S12 and 3S12 combined). As shown in Table 3, the estimated frequency was not affected by the number of colearners or its interaction with colearner questioning frequency (both F s < 1).

Consistency Between Real-time Responses and Retrospective Reports

As in Experiment 1, correlations between the number of confusion responses during the presentations and the respective postexperimental rating formats were computed within-groups and averaged across conditions using r to Z transformations. Correlations between responses during the first message and subsequent ratings of that interval were all moderately high and statistically significant (Likert = .52, self-ratings = .49, percentage of time

confused = .67; all p s $< .0001$, with 70 dfs), and they were consistently higher than correlations of responses during the first message and ratings of the second message interval (Likert = .22, self-ratings = .03, percentage of time confused = .47). Likewise, correlations between the frequency of responses during the second message and ratings of confusion during that interval were high and significant (Likert = .52, self-ratings = .32, percentage of time confused = .52; all p s $< .01$), and with the exception of the self-ratings, somewhat higher than correlations between responses during the second message and ratings of the first message interval (Likert = .39, self-ratings = .38, percentage of time confused = .41; all p s $< .01$). Once again, therefore, there is evidence that participants' summary judgments made retrospectively are consistent with monitoring judgments made during the presentations, although the differentiation between parallel and nonparallel responses and subsequent ratings is not nearly as marked as it was in Experiment 1.

Results and Discussion

Replication Analysis

Change from baseline confusion responses in the nonsocial control condition and social conditions S0 and S12 were used to test for replication. The mean change scores are plotted as a function of the frequency of colearner questions in Figure 2. The means are in the expected order: S0 = -2.0 ($SD = 5.0$), nonsocial control = -0.30 ($SD = 4.5$), and S12 = 1.3 ($SD = 2.7$). The values are also very similar to those in Experiment 1 (also shown in Figure 2). To test for replication, change from baseline scores for conditions S0 and S12 were analyzed with a 2 (frequency of colearner questions) \times 2 (order of message presentation) \times 2 (experimenter) ANOVA. The frequency of colearner questions main effect was significant, $F(1, 32) = 7.71$, $p < .01$, $MSE = 14.13$, which corroborates the previous finding. Thus, there is evidence that a colearner who asks several questions raises more doubts about participants' comprehension than does a colearner who asks no questions.

Table 3
Effects of Questioning Frequency and Number of Colearners (Experiment 3)

No. of Colearners	Colearner question frequency		ANOVA effects (F values)		
	0	12	Frequency (F)	Colearners (C)	F \times C
	Estimated frequency				
1	0.2	9.3	147.87***	< 1	< 1
3	0.1	10.3			
	Confusion responses ^a				
1	-2.0	1.3	6.31*	< 1	4.38*
3	0.4	0.7			
	Awareness				
1	1.9	5.3	53.61***	1.81	1.58
3	0.4	5.2			
	Inferred confusion				
1	2.4	5.6	33.93***	1.84	< 1
3	3.1	6.1			

Note. ANOVA = analysis of variance.

^a Change from baseline.

* $p < .05$. *** $p < .0001$.

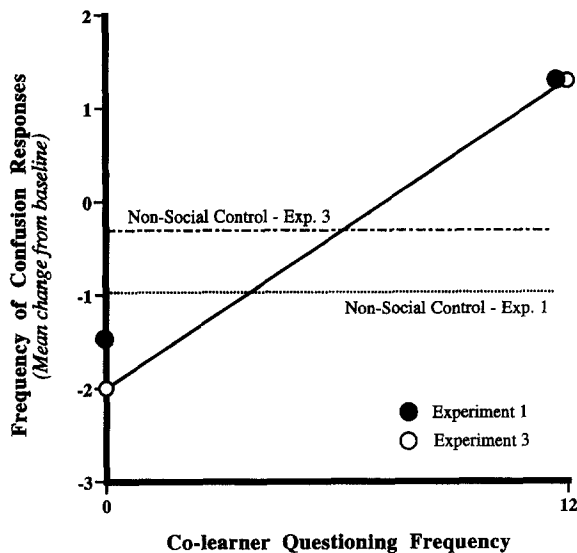


Figure 2. Mean number of participant confusion responses (change from baseline to testing) as a function of colearner questioning frequency conditions in Experiment 3. Data from identical conditions in Experiment 1 are plotted for comparison purposes.

A subsequent ANOVA that included the nonsocial control was used to test for differences between that condition and S0 and S12. A significant condition main effect, $F(2, 48) = 3.88, p < .03, MSE = 14.03$, and Dunnett directional tests ($\alpha = .05$) found no statistically significant difference between each of the two social conditions and the nonsocial control. Therefore, consistent with Experiment 1, there is no statistical evidence that the colearners who have no questions increase judged comprehension compared with the absence of social comparison information. Unlike Experiment 1, however, there is also no statistical evidence that, compared with nonsocial conditions, learners judge themselves more confused when a colearner has several questions.

Effect of the Number of Colearners

Table 3 presents the mean changes from baseline as a function of frequency of colearner questions and number of colearners. A 2 (frequency of colearner questions) $\times 2$ (number of colearners) $\times 2$ (order of message presentation) $\times 2$ (experimenter) ANOVA yielded a significant frequency of questions main effect, $F(1, 64) = 6.31, p < .02, MSE = 10.28$. However, a significant Frequency of Questions \times Number of Colearners interaction, $F(1, 64) = 4.38, p < .05$, followed by simple effects tests, indicates that the frequency main effect is attributable primarily to the single colearner conditions—a difference of 3.3 between S0 and S12, $F(1, 64) = 10.60, p < .01$. There is no evidence that increasing the number of colearners increases the impact of question asking. To the contrary, the presence of additional colearners appears to have diluted rather than magnified the effect (a difference of .3 between S0 and S12; simple effect

$F < 1$). The following analyses of postexperimental ratings provide evidence relating to that difference.

Awareness of Colearners and Inferred Colearner Confusion

As in Experiment 1, awareness was statistically higher when colearners had 12 questions ($M = 5.2, SD = 2.7$) than when they had no questions ($M = 1.2, SD = 2.2$), $F(1, 64) = 53.61, p < .001, MSE = 5.92$. However, the Frequency of Colearner Questions \times Number of Colearners interaction was not significant ($F = 1.58$), indicating that the difference in awareness levels did not depend on the number of colearners. An analysis of inferred colearner confusion yielded virtually identical results. Participants rated colearners with 12 questions to be significantly more confused ($M = 5.8, SD = 2.0$) than colearners who signaled no questions ($M = 2.8, SD = 2.7$), $F(1, 64) = 33.93, p < .0001, MSE = 4.81$, and the Frequency of Colearner Questions \times Number of Colearners interaction was not significant ($F < 1$), indicating that the difference in inferred confusion as a function of colearner questioning frequency was the same for three as for one colearner. Therefore, there was no frequency of colearner questioning effect on monitoring with multiple colearners despite evidence of differences in both awareness and inferred colearner confusion.

Retrospective Self-Reported Confusion

Retrospective reports of participant confusion were again analyzed to determine their sensitivity to condition effects. Frequency of (colearner questions by number of colearners by order of message presentation by experimenter) ANOVAs ($2 \times 2 \times 2 \times 2$) of retrospective reports yielded no significant main or interaction effects using either Likert scale, self-rating, or percentage of confused response formats (all F s < 1.5). Thus, as in Experiment 1, retrospective reports failed to reflect the effects of colearner questioning frequency that were evident in real-time responses.

In addition to replicating the findings of Experiment 1, evidence from the present experiment suggests that the same number of questions dispersed among several colearners has no greater impact than when a single colearner has questions. In fact, multiple colearners may have less impact than a single colearner. One possibility is that participants were responding to signaled questions on the basis of the number per individual rather than all colearners combined and that the impact was smaller because the number of questions per individual in the multiple colearner condition was substantially lower (4) than in the single colearner condition (12). If that were the case one might expect lower ratings of confusion in condition S12 than in condition S0. Although the ratings do not support that interpretation—as shown in Table 3, inferred colearner confusion in the multiple colearner condition was comparable ($M = 6.1$) with the level in the single colearner condition ($M = 5.6$)—this may be a function of the manner in which those ratings were obtained. For comparability between single and multiple

colearner conditions, participants were asked for judgments of the group of colearners rather than of an individual colearner. Thus, ratings of multiple colearners probably reflected their combined level of confusion rather than that of each individual, and the ratings may not provide a definitive test of this interpretation.

Experiment 4

The final experiment tested (a) whether the smaller impact of colearner questions with multiple colearner questions found in Experiment 3 would replicate and (b) whether increasing the number of questions per colearner, and thus the total number of colearner questions, would compensate for having more colearners. To accomplish this, the design included the multiple colearner conditions from Experiment 3 (3S0 and 3S12) and added a condition in which colearners indicated having substantially more questions (24; denoted as 3S24). Although the number of questions per colearner was not greater than the 12 used in the single colearner condition, it represented double the number per colearner used in Experiment 3 (i.e., 8 vs. 4) and was expected to create a detectable difference in participants' own judged level of confusion, compared with the condition in which multiple colearners asked no questions.

Method

Participants, Assignment to Conditions, and Procedure

Participants (21 men and 39 women) were from the same population and volunteered under the same conditions as those in the previous three studies. In order of arrival, they were placed by the experimenter into conditions by means of a within-gender block randomization procedure until 10 were assigned to each of the three conditions and two orders of message presentation. The stimulus materials and procedure were identical to those used in the three colearner conditions in Experiment 3.

Estimated Colearner Questioning Frequency

As in the previous experiments, all participants correctly identified the conditions in place during the first and second messages. Also similar to previous studies, participants' estimate of their colearners' questioning frequency was clearly differentiated. As shown in Table 4, the mean estimated colearner question frequencies were in the expected order, and although lower in magnitude, the estimates are in direct proportion to the difference in frequencies between conditions. Because there was no variance in condition 3S0 (all participants indicated that colearners had no questions) it was not included in the statistical analysis, which consisted of a 2 (Frequency of Colearner Questions) \times 2 (Order of Message Presentation) ANOVA. Conditions 3S12 and 3S24 differed statistically, $F(1, 36) = 7.90, p < .01, MSE = 51.84$.

Results and Discussion

Mean participant confusion responses (changes from baseline) as a function of frequency of colearner questions

Table 4

Effects of Frequency of Multiple Colearner Question Asking (Experiment 4)

Dimension	Colearner question frequency		
	0	12	24
Estimated frequency	0.0	7.6 _a	14.4 _b
Confusion responses ^a	-1.2 _a	-0.4 _{a,b}	1.9 _b
Colearner confusion	3.5 _a	4.2 _a	5.6 _b
Awareness	1.4 _a	4.8 _b	5.8 _b

Note. Noncommon letter subscripts within rows indicate means that are statistically different.

^a Change from baseline.

are shown in Table 4. A 3 (frequency of colearner questions) \times 2 (order of message presentation) ANOVA and directional Dunnett tests ($\alpha = .05$) were used to determine whether the means of conditions 3S12 and 3S24 were higher than that of condition 3S0. The frequency main effect was statistically significant, $F(2, 54) = 3.88, p < .03, MSE = 4.87$. Consistent with the results of Experiment 3, condition 3S12 was not statistically higher than condition 3S0. This indicates once again that the impact of questioning is reduced when the same number of questions (and thus a smaller number per colearner) emanate from several versus a single colearner. The difference between 3S0 and 3S24 was significant, however. Interestingly, the difference of 3.1 between the means of these conditions in the present study is approximately equivalent to that between S0 and S12 in Experiment 1 (2.8) and Experiment 2 (3.3). Thus, doubling the number of questions per colearner approximately compensated for tripling the number of colearners.

Consistent with results of the previous experiments, the frequency of participants' confusion responses paralleled their inferred level of colearner confusion. According to a 3 (frequency of colearner questions) \times 2 (order of message presentation) ANOVA on postexperimental ratings of colearner confusion, there was a significant main effect of the frequency of colearner questions, $F(2, 54) = 5.12, p < .01, MSE = 3.91$. There was a significant difference between conditions 3S0 and 3S24; however, unlike in Experiment 3 there was only a small and nonsignificant difference in inferred colearner confusion between 3S0 and 3S12 (all $ps < .05$). In the present study, therefore, the levels of inferred colearner confusion approximately paralleled participants' own confusion responses. A similar analysis conducted on postexperimental ratings of participants' awareness of colearners indicated that several colearner questions appear sufficient to have made participants aware of colearners, but that double the number had relatively little additional impact.

General Discussion

The results support the hypothesized influence of colearner questioning on comprehension monitoring. Consistent across three experiments, learners' awareness of their

colearners' questions about material they were studying affected judgments of their own level of comprehension. The more questions colearners signaled having, the more confused learners admitted they were. There is also evidence that, compared with conditions in which no social comparison information was available, colearner questions increased self-judged confusion. Therefore, as proposed by Brown and Palincsar (1988), the presence of other learners, for example in collaborative work groups, creates conditions for the externalization of cognitive dissatisfaction.

There was, however, no statistical justification for the obverse—that persons judge themselves less confused in the absence of questioning by others than they would have without social comparison information (the externalization of cognitive satisfaction). Nevertheless, the order of the means was consistent in two experiments. Although it was noted that range restriction (a floor effect) is unlikely, it is possible that participants were less aware of colearners who asked no questions (as indicated by postexperimental ratings), thereby decreasing colearners' social impact and attenuating the effect of this condition on participants' self-judged confusion.

If awareness of nonquestioning colearners was a factor in determining their impact, then we would expect participants who were more aware of colearners to have been less confused than those who were less aware. An internal analysis provided support for this hypothesis. When participants in condition S0 from Experiments 1 and 3 were combined and partitioned into those that were higher versus lower in awareness, those who indicated being at least moderately aware of their colearners (ratings ≥ 2 on a 0 to 9 point scale, $n = 13$) had a mean change from baseline of -3.9 ($SD = 6.0$), whereas the change from baseline of those totally unaware of colearners (rating of 0, $n = 18$) was only -1.7 ($SD = 2.7$), $F(1, 27) = 4.83$, $p < .05$, $MSE = 13.06$. Although tentative because the association between awareness and judged confusion is based on correlational rather than experimental evidence, it suggests that salient colearners who ask no questions may indeed have the effect of reducing self-judged confusion. Having colearners actually present, or at least in visual contact, rather than unseen in another physical location, may be required to create the level of awareness necessary to manifest the effects of colearners not asking questions. Such conditions would be closer approximations to actual classroom settings and collaborative learning arrangements.

Participant differences in retrospectively reported confusion and real-time monitoring judgments were related. That is, participants having more confusion episodes while they were viewing the messages were more likely to report having been confused. Nevertheless, retrospective ratings were apparently not sensitive to the effects of colearner questioning conditions, which is consistent with previous studies of comprehension monitoring (e.g., Markman, 1985; see also Jacobs & Paris, 1987; Nisbett & Wilson, 1977). This has important implications for the study of social effects on monitoring in interactive settings. If retrospective reports were incapable of detecting the effects in the controlled environment of the present study, it is highly un-

likely they would do so given the greater complexity of uncontrolled interactive settings. Some form of on-line assessment may be required (Pressley & Ghatala, 1990).⁴

The link from others' questioning (or lack thereof) to self-monitoring appears to be through inferred levels of colearner confusion. Because postexperimental ratings revealed that inferred colearner intelligence was not affected by questioning frequency, it is possible to conclude that participants attributed the source of colearner confusion to the stimulus material that they were attempting to comprehend (i.e., a stimulus attribution). From the participants' perspective, "others are finding it difficult (or easy) to understand the speech, therefore the material is (not) confusing and I, too, may be more (or less) confused." It follows that learning settings that facilitate such task attributions would produce greater influences on monitoring, whereas conditions that result in attributions to sources other than task difficulty (e.g., persons) would decrease the effect. In classrooms, the degree of perceived teacher support for student questions would be one such factor (Karabenick & Sharma, 1994). According to the augmentation principle (Kelley, 1967; Weiner, 1985; Weiner, Graham, Taylor, & Meyer, 1983), colearners who ask questions with nonsupportive teachers would be perceived as having done so in spite of normative prohibitions, leading to the inference that they must have really needed to ask (i.e., been very confused). According to the principle of discounting, the reverse should be true as well; the value of questions as indicators of colearner confusion would be reduced when teachers encourage questions, an additional cause to which questioning can be attributed. Students' history of questioning represents a second source of information that may increase or decrease confusion as the source of questioning. Colearner questioning would more likely be attributed to information-induced confusion if that person tended not to ask many questions. Colearner who ask questions frequently may be perceived as more confused, but with that state attributed to internal causes (e.g., lack of ability or effort) rather than to the material they are trying to comprehend.

Whether college students would be aware of these effects was explored by presenting them with a series of classroom scenarios that factorially combined levels of teacher support and prior colearner questioning frequency (Karabenick,

⁴ One reason for the disparity could be that retrospective reports asked participants to judge their degree of confusion during two time periods—baseline and testing. Whereas participants who were generally more confused retrospectively reported having been so, as evidenced by the within-group correlations, the differences between the retrospective ratings of confusion during baseline and testing did not reflect condition effects. This is not to suggest that retrospective reports are invalid, but rather that they may not substitute for real-time assessment (see Baker, 1989, for a review of measurement issues). It is possible that retrospective reports obtained after each time period might have improved their quality; however, real-time information was of primary concern rather than a comparison between measures, and interrupting the procedure to obtain retrospective reports following baseline was deemed too obtrusive.

1993). As predicted, students reported they would be more confused when colearners asked questions under conditions in which teachers discouraged rather than encouraged questions. Students also reported they would be more confused when the questioner had previously asked few rather than many questions. Such results warrant further examination to determine whether these effects are detectable in real time.

The importance of understanding the effects of colearner questioning on monitoring takes on greater significance with the increasing prevalence of classroom technology, such as desktop keypads, that can provide instructors with real-time information about student comprehension. Not only does the technology have considerable implications for teachers (e.g., indicating when slowing down or repeating information may be necessary), but it could also alter the metacognitive environment for students if such information were to be publicly displayed. With this technology, it would not even be necessary for colearners to ask questions from which to infer confusion; rather, student confusion or perplexity could be obtained and displayed directly. Cumulative displays that provide for anonymity would even remove some of the threat attendant upon such responses (Karabenick & Knapp, 1988, 1991). It is suggested that the effects on monitoring found in the present study be taken into consideration when designing how, or if, student response information should be displayed, as well as determining its impact on student monitoring and learning.

Although the present study detected effects of colearner questioning on judged comprehension, several limitations are noted. First, the study was not designed to detect changes in actual comprehension. Although participants completed a brief test following each presentation, that procedure was used to enhance the experiment's cover story rather than to determine how much persons had actually learned from the material presented to them. Nor did I assess whether the effects of colearner questions on comprehension monitoring triggered more attempts at remediation (e.g., self-questioning; King, 1992). I suggest that additional studies are needed to determine the self-regulatory consequences of colearner questions. Also noted is that the degree of experimental control that was gained by using unseen (and simulated) colearners may have underestimated the judged comprehension consequences of the absence of questions. Stronger effects of having actual colearners present may outweigh the additional complexity (e.g., gender, status, and other differences between participants and colearners) that would have to be taken into consideration with greater realism. Moving toward the study of more natural settings would also introduce a host of additional factors that could affect comprehension judgments, such as the content of questions and of answers to those questions. How would judged confusion be affected when, for example, a colearner's question raises issues that other students had not considered? Or when an instructor provides an inadequate answer to the student's question?

Experimental evidence of socially mediated information on comprehension monitoring is important given the extent to which learning occurs in social contexts, or, as some would argue (Kozulin & Presseisen, 1995; Vygotsky,

1978), is a function of social interaction. Whether in small study groups or large classrooms, in addition to their contents, the knowledge that others are perplexed by information they are receiving, and inferences about why they are confused, can apparently trigger cognitive dissatisfaction, whereas the absence of questions can lull learners into a false state of cognitive satisfaction. We would do well to understand this process more completely and determine the social conditions that moderate it. This phenomenon should also be added to the ways that social and cognitive processes interact (Levine et al., 1993; McGivern, Levin, Pressley, & Ghatala, 1990).

References

- Alexander, P. A. (1995). Superimposing a situation-specific and domain-specific perspective on an account of self-regulated learning. *Educational Psychologist*, 30, 189-193.
- Asch, S. E. (1955). Opinions and social pressure. *Scientific American*, 19, 31-35.
- Baker, L. (1989). Metacognition, comprehension monitoring, and the adult reader. *Educational Psychology Review*, 1, 3-38.
- Baker, L., & Anderson, R. I. (1982). Effects of inconsistent information on text processing: Evidence for comprehension monitoring. *Reading Research Quarterly*, 17, 281-294.
- Bender, T. A. (1993, April). Predicting postfeedback performance from students' confidence in their responses. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.
- Bond, C. F., Jr., & Titus, L. J. (1983). Social facilitation: A meta-analysis of 241 studies. *Psychological Bulletin*, 94, 265-292.
- Brown, A. L., & Palinscar, A. S. (1988). Guided, cooperative learning and individual knowledge acquisition. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 393-451). Hillsdale, NJ: Erlbaum.
- Cottrell, N., Eisenberger, R., Sekerak, G. J., & Rittle, R. (1968). Social facilitation of dominant responses by the presence of an audience and the mere presence of others. *Journal of Personality and Social Psychology*, 51, 245-250.
- Darling, A. L. (1989). Signaling non-comprehensions in the classroom: Toward a descriptive topology. *Communication Education*, 38, 34-40.
- Dillon, J. T. (1986). Student questions and individual learning. *Educational Theory*, 36, 333-341.
- Epstein, W., Glenberg, A. M., & Bradley, M. M. (1984). Coactivation and comprehension: Contribution of text variables to the illusion of knowing. *Memory & Cognition*, 12, 355-360.
- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7, 117-140.
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new era in cognitive-developmental inquiry. *American Psychologist*, 34, 906-911.
- Flavell, J. H. (1987). Speculations about the nature and development of metacognition. In F. Weinert & R. Kluwe (Eds.), *Metacognition, motivation, and understanding* (pp. 21-29). Hillsdale, NJ: Erlbaum.
- Garner, R. (1981). Monitoring of passage inconsistency among poor comprehenders: A preliminary test of the "piecemeal processing" explanation. *Journal of Educational Research*, 74, 159-165.
- Glaser, R., & Bassok, M. (1989). Learning theory and the study of instruction. *Annual Review of Psychology*, 40, 631-666.
- Goethals, G., & Darley, J. M. (1987). Social comparison theory:

- Self-evaluation and group life. In B. Mullen & G. Goethals (Eds.), *Theories of group behavior* (pp. 21–48). New York: Springer-Verlag.
- Grabe, M., Antes, J., Thorson, I., & Hahn, H. (1987, April). *The impact of questions and instructions on comprehension monitoring*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.
- Jackson, M. M. (1987). Social impact theory: A social forces model of influence. In B. Mullen & G. R. Goethals (Eds.), *Theories of group behavior* (pp. 111–124). New York: Springer-Verlag.
- Jacobs, J. H., & Paris, S. G. (1987). Children's metacognition about reading: Issues in definition, measurement, and instruction. *Educational Psychologist*, 22, 255–278.
- Karabenick, S. A. (1993, June). *Co-learner questions affect comprehension monitoring and questioning intentions*. Presented at the annual convention of the American Psychological Society, Chicago, IL.
- Karabenick, S. A., & Knapp, J. R. (1988). Effects of computer privacy on help-seeking. *Journal of Applied Social Psychology*, 18, 461–472.
- Karabenick, S. A., & Knapp, J. R. (1991). Relationship of academic help seeking to the use of learning strategies and other instrumental achievement behavior in college students. *Journal of Educational Psychology*, 83, 221–230.
- Karabenick, S. A., & Sharma, R. (1994). Perceived teacher support of student questioning in the college classroom: Its relation to student characteristics and role in the classroom questioning process. *Journal of Educational Psychology*, 86, 90–103.
- Kelley, H. (1967). Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska Symposium on Motivation* (pp. 192–238). Lincoln: University of Nebraska Press.
- King, A. (1992). Comparison of self-questioning, summarizing, and notetaking-review as strategies for learning from lectures. *American Educational Research Journal*, 29, 303–323.
- Kozulin, A., & Presseisen, B. Z. (1995). Mediated learning experience and psychological tools: Vygotsky's and Feuerstein's perspectives in a study of student learning. *Educational Psychologist*, 30, 67–75.
- Kulhavy, R. W., & Stock, W. A. (1989). Feedback in written instruction: The place of response certitude. *Educational Psychology Review*, 1, 279–308.
- Latané, B. (1981). The psychology of social impact. *American Psychologist*, 36, 343–356.
- Latané, B., & Darley, J. M. (1970). *The unresponsive bystander: Why doesn't he help?* New York: Appleton-Century-Crofts.
- Latané, B., & Wolf, S. (1981). The social impact of majorities and minorities. *Psychological Review*, 88, 438–453.
- Levin, J. R., Serlin, R. C., & Seaman, M. A. (1994). A controlled, powerful multiple-comparison strategy for several situations. *Psychological Bulletin*, 115, 153–159.
- Levine, J. M., Resnick, L. B., & Higgins, E. T. (1993). Social foundations of cognition. *Annual Review of Psychology*, 44, 585–612.
- Markman, E. M. (1985). Comprehension monitoring: Developmental and educational issues. In S. F., Chipman, J. W. Segal, & R. Glaser (Eds.), *Thinking and learning skills* (Vol. 2, pp. 275–291). Hillsdale, NJ: Erlbaum.
- Markman, E. M., & Gorin, L. (1981). Children's ability to adjust their standards for evaluating comprehension. *Journal of Educational Psychology*, 73, 320–325.
- McGivern, J. E., Levin, J. R., Pressley, M., & Ghatala, E. S. (1990). A developmental study of memory monitoring and strategy selection. *Contemporary Educational Psychology*, 15, 103–115.
- Miller, D. T., & Prentice, D. A. (1994). Collective errors and errors about the collective. *Personality and Social Psychology Bulletin*, 20, 541–550.
- Morris, W. N., & Miller, R. S. (1975). The effects of consensus-breaking and consensus-preempting partners on reduction of conformity. *Journal of Personality and Social Psychology*, 11, 215–223.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231–259.
- Paris, S. G., Lipson, M. Y., & Wixon, K. K. (1983). Becoming a strategic reader. *Contemporary Educational Psychology*, 8, 293–316.
- Pintrich, P. R. (1989). The dynamic interplay of student motivation and cognition in the college classroom. In M. Maehr & C. Ames (Eds.), *Advances in motivation and achievement: Motivation enhancing environments* (Vol. 6, pp. 117–160). New York: JAI Press.
- Pressley, M., & Ghatala, E. S. (1990). Self-regulated learning: Monitoring learning from text. *Educational Psychologist*, 25, 19–33.
- Rosenthal, R., & Rosnow, R. L. (1985). *Contrast analysis: Focused comparisons in the analysis of variance*. New York: Cambridge University Press.
- Rosnow, R. L., & Rosenthal, R. (1988). Focused tests of significance and effect size estimation in counseling psychology. *Journal of Counseling Psychology*, 35, 203–208.
- Ruble, D. (1983). The development of social comparison processes and their role in achievement-related self-socialization. In E. T. Higgins, D. N. Ruble, & W. W. Hartup (Eds.), *Social cognition and social development: A sociocultural perspective* (pp. 134–157). New York: Cambridge University Press.
- Ruble, D., & Frey, K. (1991). Changing patterns of comparative behavior as skills are acquired: A functional model of self-evaluation. In J. Suls & T. Wills (Eds.), *Social comparison: Contemporary theory and research* (pp. 79–113). Hillsdale, NJ: Erlbaum.
- Schunk, D. H., & Zimmerman, B. J. (Eds.). (1994). *Self-regulation of learning and performance: Issues and educational applications*. Hillsdale, NJ: Erlbaum.
- Sherif, M. (1935). A study of some social factors in perception. *Archives of Psychology*, No. 187.
- Suls, J. M., & Wills, T. A. (Eds.). (1991). *Social comparison: Contemporary theory and research*. Hillsdale, NJ: Erlbaum.
- Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*. (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). Cambridge, MA: Harvard University Press.
- Weiner, B. (1979). A theory of motivation for some classroom experiences. *Journal of Educational Psychology*, 71, 3–25.
- Weiner, B. (1985). An attribution theory of achievement motivation and emotion. *Psychological Review*, 92, 548–573.
- Weiner, B., Graham, S., Taylor, S., & Meyer, W. U. (1983). Social cognition in the classroom. *Educational Psychologist*, 18, 109–124.
- Weinstein, C. E., & Mayer, R. E. (1985). The teaching of learning strategies. M. Wittrock (Ed.), *Handbook of research on teaching* (pp. 315–327). New York: Macmillan.
- Zajonc, R. B. (1965). Social facilitation. *Science*, 149, 269–274.
- Zimmerman, B. J. (1992). Self-regulated learning and academic achievement: An overview. *Educational Psychologist*, 25, 3–18.
- Zimmerman, B. J. (1995, April). *Goal setting and self-monitoring*

in the self-regulated learning of motoric skill. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Zimmerman, B. J., Greenberg, D., & Weinstein, C. E. (1994). Self-regulating academic study time: A strategic approach. In D. H. Schunk & B. J. Zimmerman (Eds.), *Self-regulation of learning and performance: Issues and Educational Applications* (pp. 181–199). Hillsdale, NJ: Erlbaum.

Zimmerman, B. J., & Martinez-Ponz, M. (1992). Perceptions of

efficacy and strategy use in the self-regulation of learning. In D. H. Schunk & J. Meece (Eds.), *Student perceptions in the classroom: Causes and consequences*. (pp. 185–207). Hillsdale, NJ: Erlbaum.

Received June 28, 1995

Revision received April 8, 1996

Accepted April 15, 1996 ■

New Editors Appointed, 1998–2003

The Publications and Communications Board of the American Psychological Association announces the appointment of five new editors for 6-year terms beginning in 1998.

As of January 1, 1997, manuscripts should be directed as follows:

- For the *Journal of Experimental Psychology: Animal Behavior Processes*, submit manuscripts to Mark E. Bouton, PhD, Department of Psychology, University of Vermont, Burlington, VT 05405-0134.
- For the *Journal of Family Psychology*, submit manuscripts to Ross D. Parke, PhD, Department of Psychology and Center for Family Studies-075, 1419 Life Sciences, University of California, Riverside, CA 92521-0426.
- For the Personality Processes and Individual Differences section of the *Journal of Personality and Social Psychology*, submit manuscripts to Ed Diener, PhD, Department of Psychology, University of Illinois, 603 East Daniel, Champaign, IL 61820.
- For *Psychological Assessment*, submit manuscripts to Stephen N. Haynes, PhD, Department of Psychology, University of Hawaii, 2430 Campus Road, Honolulu, HI 96822.
- For *Psychology and Aging*, submit manuscripts to Leah L. Light, PhD, Pitzer College, 1050 North Mills Avenue, Claremont, CA 91711-6110.

Manuscript submission patterns make the precise date of completion of the 1997 volumes uncertain. Current editors, Stewart H. Hulse, PhD; Ronald F. Levant, EdD; Russell G. Geen, PhD; James N. Butcher, PhD; and Timothy A. Salthouse, PhD, respectively, will receive and consider manuscripts until December 31, 1996. Should 1997 volumes be completed before that date, manuscripts will be redirected to the new editors for consideration in 1998 volumes.