Journal of Personality and Social Psychology
1996, Vol. 71, No. 5, 915–927

Copyright 1996 by the American Psychological Association, Inc.
0022-3514/96/$3.00

The Forgotten Variable in Conformity Research: Impact of Task Importance on Social Influence

Robert S. Baron, Joseph A. Vandello, and Bethany Brunsman
University of Iowa

Two studies examined how incentives for accuracy (task importance) affected the social influence of inaccurate confederates in a modified Asch situation (S. E. Asch, 1951). Not unexpectedly, when task difficulty was low, incentives for accuracy reduced the social impact of (inaccurate) confederates (Study 1). However, when task difficulty was increased, the reverse was true, with individuals conforming more to an inaccurate confederate norm when incentives for accuracy were high (Studies 1 and 2). The results are discussed in terms of possible mediating mechanisms and also in terms of their historical and pragmatic implications.

Social psychologists have long been interested in social influence and conformity. The body of experimental research in the area has spanned more than half a century, dating back to the Sherif's (1935) pioneering research on the autokinetic effect. Over this time, researchers have established that a wide range of personality and situational variables affect conformity, including group size (e.g., Asch, 1956; Wilder, 1977), fear (Darley, 1966), unanimity (e.g., Asch, 1956), ethnicity (e.g., Malof & Lott, 1962), group cohesiveness (e.g., Back, 1951), status in group (e.g., Berkowitz & Macaulay, 1961), and judgment difficulty (e.g., Deutsch & Gerard, 1955), to name a few. Despite the intense research attention focused on variables that affect conformity and social influence, one variable has been largely neglected. This variable, which we refer to below as *task importance,* concerns the extent to which making correct or accurate judgments mediates important rewards and punishments for research participants. Stated differently, this variable corresponds to how much pressure participants feel to offer accurate judgments on the task in question. At present very little is known about how this variable affects susceptibility to social influence.

In most conformity studies, participants are seldom offered any direct incentive for accuracy (or threatened with punishment for error) and, although a reasonable level of motivation to be accurate can be inferred (given the laboratory setting), this motivation is seldom measured and hardly ever manipulated. Indeed, despite the 60 years of research on this topic since Sherif (1935), we found only three conformity studies that manipulated task importance, and these studies yielded contradictory outcomes. Di Vesta (1959) manipulated motivation by telling half of his participants that their responses correlated highly with intelligence. These highly motivated participants conformed less (on perceptual and attitude judgments) than less motivated participants. In another study, by Crowne and Liverant (1963), female participants (but not male participants) conformed more when they bet on the accuracy of their responses (participants could win up to $3). A third report, by Crutchfield (cited in Krech, Crutchfield, & Ballachey, 1962, pp. 220–221) found that the offer of $10 per person for being a member of the most accurate 5-person group (out of 12 groups) increased conformity across a mixed set of factual, perceptual, and attitude judgments in one substudy. This incentive effect failed to replicate in a second substudy, however, when payment was based on each individual's (vs. the group's) performance. In the present study we extended this research, tested a conceptual explanation for the effects of task importance on social influence, and offer a tentative account for the discrepancies in the studies just discussed.

The general issue of how task importance affects social influence has several interesting implications. At an empirical level, the lack of data regarding the variable of task importance raises serious questions of generality. One could dismiss the conformity effect as a laboratory "hothouse" phenomenon that occurs because the potential face-to-face rejection of peers is far more important to participants than their accuracy on some unimportant "scientific" test of perception or social judgment. Indeed, undergraduate students are not loath to offer such criticisms when presented with the classic conformity studies of Asch (1951, 1956), Sherif (1935), Deutsch and Gerard (1955), and others. From this perspective, social influence effects would vanish on important decisions that have meaning-

Robert S. Baron, Joseph A. Vandello, and Bethany Brunsman, Department of Psychology, University of Iowa. Joseph A. Vandello is now at the Department of Psychology, University of Illinois at Urbana–Champaign.

We thank Pam Young and Bev Hamann for their invaluable secretarial assistance. Studies 1 and 2 were supported in part by funds from the University of Iowa Undergraduate Honors Program. We thank Matt Brick, Amy Cox, Niki Delemont, Julie Dotson, Susan Lui, Michael Marchman, Shannon Mitchell, David Peterson, and Jeff Vittengle for help as experimenters and confederates in Study 1 and Nandita Chowla, Kelly Chaplin, Shelly Dorman, Kory Frederick, Debra Geier, Larissa Johnston, Susan Kruse, Kevin McKay, Karen Mazaika, Stuart Miller, Dan Verbeten, Andy Weigel, and Amy Zeitz for similar help in Study 2.

Correspondence concerning this article should be addressed to Robert S. Baron, Department of Psychology, University of Iowa, Iowa City, Iowa 52242.

ful consequences for the individual. In one sense, this is an optimistic view that argues that people will (a) act rationally and (b) be impervious to inaccurate majority pressure when "the chips are down." Without more data regarding the impact of task importance on social influence, it is hard to either rebut or evaluate the validity of such criticisms, but recent failures to replicate classic social influence phenomena (Nicholson, Cole, & Rocklin, 1985; Perrin & Spencer, 1981) do little to quiet such doubts.

The issue of whether and how task importance affects social influence also has important practical implications. In contrast to the optimistic view offered above, numerous scholars have suggested that social influence will have a strong impact on decisions of substantial importance (e.g., Le Bon, 1895/1960; Sherif, 1935; Hoffer, 1951; Janis, 1972). Indeed, various writers have alluded to the importance of social influence in explaining the success of the Nazi party in Germany in the 1930s and 40s and other totalitarianism (and demagogic) movements as well (e.g., Hoffer, 1951; Shirer, 1960; Weightsman, 1984). In a similar view, Janis (1972) contended that crisis-laden political/scientific/technical/military decisions of great importance were particularly likely to trigger the various conformity mechanisms thought to underlie groupthink phenomena. Janis's notion that stress exacerbates social influence is quite relevant to our discussion above. That is, it seems safe to assume that one potent consequence of increasing the threatening consequences of an incorrect decision is to heighten the (task) importance of that decision (i.e., the desire for accurate decisions on the part of the decisionmaker).[1] Indeed, one reason Janis's (1972) groupthink perspective has proven so provocative is that it predicts that as the importance (or threat level) of a judgment increases, so does the impact of social influence, even in cases where in retrospect that social influence is wrong, irrational, emotional, or based on stereotypes or flawed problem solving. This prediction is intriguing given that, commonsensically, it would seem that one would be *less* likely to succumb to such inaccurate social influence as the incentive for accuracy increases. This is obviously a prediction with far-reaching political and social implications. Therefore, it is surprising that it has received so little direct research attention.

At a theoretical level, several explanations can be offered to account for this hypothesized relationship between task importance and social influence. First, when judgment outcomes are important, participants may be more likely to engage in social comparison as part of their epistemological search for an accurate and informed answer (e.g., Kruglanski, 1989). Second, to the extent that task importance is associated with stress, feelings of dependency, low self-esteem, or needs for affiliation, it may make individuals more susceptible to social influence (Darley, 1966). A closely related idea is that stress and emotion are thought to tax attentional capacity (cf. S. Cohen, 1978), thereby increasing the tendency of individuals to rely on various cognitive and social heuristics such as social consensus and normative pressure when making a judgment (Baron, 1986). These positions all point to a simple positive relationship between task importance and accepting social influence. A more complex view, however, (and one we prefer) is that the effect of task importance on social influence will be moderated by the difficulty level of the judgment in question, with high task importance lowering social influence on judgments of low difficulty and increasing social influence on judgments of high difficulty. These predictions derive from the assumption (Deutsch & Gerard, 1955; Jones & Gerard, 1967) that individuals exposed to social influence pressure (i.e., disagreement from others) are subjected to conflicting pressures; several forces encourage them to resist social influence (self-trust, desire for independence, etc.) while others encourage them to accept influence (trust in the shared group opinion, i.e., informational social influence, and group power, i.e., normative social influence). Varying the strength of these forces is thought to affect the degree to which social influence is accepted by the individual, and a good deal of research supports these views (cf. Baron, Kerr, & Miller, 1992; Jones & Gerard, 1967; Levine, 1989).

Keeping in mind this "conflict" model, consider how task importance might affect social influence on judgments of extreme ease (comparable to those used in the original Asch [1951] paradigm in which alone-controls were correct on over 98% of trials). In such cases, of course, the participant is typically faced with a group agreeing on blatantly incorrect estimates on a number of trials. In this situation, the participants should have reasonable confidence in their own (correct) perceptions. Consequently, any social influence here is likely to be driven primarily by normative social influence. This, in turn, should produce public compliance with little change in participants' private perceptions regarding the judgment (Asch, 1951). In this case, increasing task importance (say by heightening rewards for accurate judgments) should provide the participant with added reason to resist any normative group pressure and stick with his or her own (confident and correct) judgment. Stated differently, it is one thing to depart from one's own confident judgment on a trivial matter but quite another to so acquiesce when there are substantial consequences at stake. This explanation captures much of the logic offered by the "optimistic view" outlined above.

In contrast, when task difficulty increases (and participants' confidence and self-trust drops), the participants cannot as easily rely on their own perceptions as a guide to accuracy. In cases of high judgment difficulty, where the participant is highly uncertain, the most sensible guide to a correct response is the responses of others, particularly if they are in unanimous agreement. This, of course, reprises Festinger's (1954) classic

[1] *Stress* is conceptualized here as stimuli or events that threaten the well-being of the organism. These may involve the presentation of aversive stimuli, the removal or denial of rewarding stimuli, ego threat, social sanction or embarrassment, uncertainty or conflict regarding important issues, and so on. It is important to note that stress will not invariably elevate task importance. Rather, it seems likely to do so on stress-relevant tasks where stress exposure is related to the quality of task judgments. Complicating this picture somewhat is the fact that it does seem likely that most judgmental tasks of high importance are a cause of some stress, because by definition performance on such tasks mediates valued rewards and punishments. Thus, although stress will often (but not always) increase task importance (depending on the stress "relevance" of the task), high task importance will generally elevate stress levels.

arguments regarding the importance of social reality under conditions of high judgment ambiguity. As a result, if high task importance increases participants' desires for accurate judgments, it should increase social influence on highly ambiguous (i.e., difficult) judgmental tasks. A shorthand view of this overall position is that heightening task importance (i.e., the participant's "stake" in being accurate) should decrease public compliance effects such as those reported by Asch (1951) but should increase those social influence effects thought to affect private internalized opinion change (e.g., Sherif, 1935).

In short, the effects of task importance on social influence should be moderated by task difficulty.[2] To the extent that the data corroborate this prediction, prior inconsistencies in the literature on task importance and conformity could be due, in part, to any differences in task difficulty between the three prior critical studies.[3] To examine these issues, we exposed research participants to inaccurate confederate responses in two studies ostensibly concerned with eyewitness identification.

Study 1

Method

Overview. To camouflage the well-known Asch (1951) conformity paradigm, we masked our procedure as a study of eyewitness accuracy that ostensibly was focusing on duration of exposure as a key variable. Participants were exposed in two tasks to various eyewitness identification slides in the presence of two confederates who gave unanimous and incorrect judgments on key trials prior to the participants' judgment. Task difficulty (high or low) was a compound manipulation that simultaneously varied time of exposure to the slides (on both tasks) and frequency of exposure to the slides (on the lineup task). We manipulated task importance (high or low) by telling participants in the high-importance condition that the test procedure (i.e., performance on both tasks) was a verified measure of eyewitness ability and that participants whose accuracy placed them in the top 12% of scores would receive $20 at the end of data collection. Participants in the low-importance condition were told that the test procedure was part of a pilot study. Two additional control groups of individuals (baseline controls) responded to the eyewitness tasks without the influence of confederates, to establish baseline error rates for the easy and difficult tasks and to verify the effectiveness of the difficulty manipulation.

Participants. Participants were 95 Introductory Psychology students at the University of Iowa (37 men and 58 women) who participated in the study as one means of fulfilling a course requirement. Of these, 5 participants were excluded because they surmised the general nature of the experiment (i.e., that it involved nonindependent judgments). Sixty of the remaining participants (23 men, 37 women) were assigned to the four experimental conditions. The remaining 30 participants (13 men, 17 women) were assigned to the baseline control conditions. In addition, 94 students from an introductory social psychology class participated as post hoc judges of the task importance manipulation.

Materials and instructions. Two eyewitness identification tasks were used: a lineup task and a description task. On both tasks the stimuli were slides of black ink drawings of male figures projected to 2 × 2.5 ft (0.6 m × 0.75 m) images on a white screen located 9 ft (approximately 3 m) in front of the participants. All drawings were full frontal body depictions (see Figure 1). On the lineup task, each trial consisted of exposure to an initial perpetrator slide followed by a lineup slide consisting of the perpetrator and three additional men. In all cases, the perpetrator was present in the lineup. Thirteen trials (i.e., 26 slides) were presented. The lineup task was introduced with the following instructions, which were read to participants by the experimenter:

> This next task will measure your ability to recognize individuals in lineups. First, I will show you a picture of a single individual; then, I will show a lineup of four individuals. The previous individual will be one of the four, but he may be dressed differently the second time. Your job is to call out which person is the same as the original individual. The pictures will be shown for varying lengths of time. I want you to call out numbers between 1 and 4 when answering; the person on the left being person 1 and continuing right through person four.

Despite these instructions, stimuli on both tasks were exposed at one fixed duration at each level of task difficulty. On the lineup task, participants were simply asked to indicate which individual in the lineup slide had appeared previously on the perpetrator slide. Note that this occurred after exposure to (i.e., in the absence of) the lineup slide.

The seven slides used in the description task consisted of full frontal body drawings of two male figures. The description task was introduced with the following instructions, which were read to participants:

> On this final task, you will be shown slides of two individuals for varying amounts of time. When looking at the individuals, I want you to remember details of their appearances, because I will ask you questions about them after the slide is turned off.

---

[2] This moderator prediction can also be explained in terms used by message-processing theories of persuasion (Chaiken, 1987; Kruglanski, 1989; Petty & Cacioppo, 1986). When task difficulty is low, participants should respond to a task of high importance (i.e., high involvement) by attending carefully to their (clear) perceptions of the stimuli and relatively less to the social cues from their fellows. In short, in this setting, participants' heightened desire to process carefully and accurately (under high task importance) can be fulfilled by focusing on their own (adequate) perceptions of the stimuli. However, as task difficulty increases, participants eventually should lose the ability to process adequately the necessary stimuli regardless of their motivation to process carefully. Under such conditions of high ambiguity, participants' best guesses about an accurate response will often stem from information regarding social consensus. Here a strong desire for accuracy should lead to *stronger* attention to social cues from peers, because this may be the only rational means of achieving judgment accuracy. This of course would lead to greater social influence as task importance increases.

[3] Di Vesta (1959) used a mixed set of stimuli, including Asch-like line judgments as well as attitude and informational judgments. In accordance with our argument (he reported less conformity under conditions of high importance), his task seemed fairly easy, with baseline (i.e., control) error rates of less than 10% in both studies. Crowne and Liverant's (1963) study (which found greater conformity under conditions of high importance) required participants to distinguish the larger of two groups of dots after a 1-s exposure. Although the task was described as unambiguous on the basis of pilot work, exposure times were brief (1 s vs. 5 s and 10 s in the easy conditions of the present research), and participant confidence ratings on the trials averaged only 5.9 on a scale of 1 to 10, suggesting that participants were somewhat uncertain. Finally, the finding that betting increased conformity was apparent only among female participants. Men actually tended to conform less in the betting condition. In the absence of a baseline control cell it is hard to interpret these data. In Crutchfield's report (cited in Krech et al., 1962, p. 209), task details are sketchy, but apparently a variety of judgment types were used (opinions, facts, perceptions, etc.) of mixed difficulty levels.

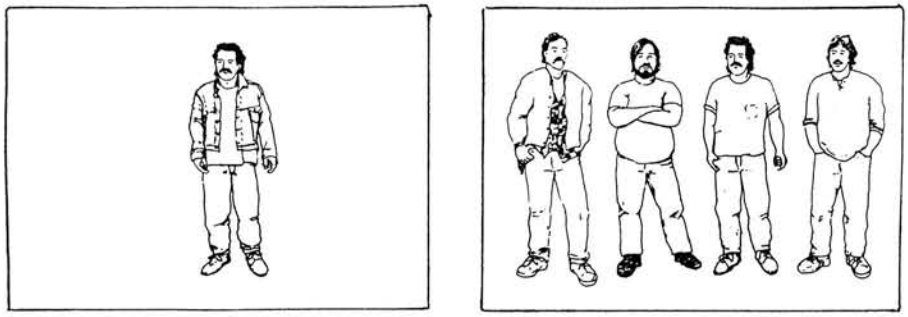

Figure 1. Sample stimulus for the lineup task.

On the description task, participants were asked (just after exposure) two (descriptive) questions per slide (e.g., "Which man was taller?" "Did the man on the left wear glasses?" "Did either man wear a wrist watch?" etc.). Thus, the description task involved responding (from memory) to a total of 14 questions. The 26 lineup slides always preceded the 7 description task slides, and all slides were shown in a fixed order to all participants.

Manipulation of task difficulty. Participants in the low-task-difficulty condition viewed (both perpetrator and lineup) slides in the lineup task for 5 s each and viewed slides in the description task for 10 s each. In addition, these individuals were shown each lineup sequence (i.e., perpetrator slide–lineup slide) twice before responding. In the high-task-difficulty condition, participants viewed lineup task slides for 0.5 s (once) and viewed description task slides for 1 s. The interstimulus interval in the lineup condition was approximately 1 s (i.e., the time it took for the slide projector to advance).

Baseline control study. We conducted a pilot study to establish a baseline error rate. Baseline control participants (n = 30) entered the experimental room in groups of 4 or 5 and were seated at individual desks facing the projection screen, which was approximately 9 ft (3 m) in front of them. On each of the desks were two sheets: a task response sheet and a questionnaire. The experimenter told the participants that they were participating in a pilot study on eyewitness testimony. They were instructed that the university was interested in developing a test to measure eyewitness accuracy, and they would be shown slides in which they were to pick individuals from a lineup (lineup task) or remember details of individuals' appearances (description task; see above).

Control participants viewed slides on the two lineup tasks under either high or low difficulty conditions (see above). After each trial, participants recorded their responses individually (and privately) on task response sheets. At the conclusion of the tasks, participants were asked to complete a questionnaire. The questionnaire contained four items measuring participants' certainty on 7-point scales (see *Dependent measures* section).

Study 1 procedure. Experimental participants in Study 1 waited outside of the experimental laboratory with two confederates (who with the experimenter were blind to all experimental hypotheses in both Studies 1 and 2). Confederate teams were either all male or mixed gender. As noted below, this did not affect conformity rates. All three were met by the experimenter, who led them into the room. The two confederates appropriated the first two chairs, leaving the chair on the far right for the naive participant. Chairs were spaced 2/3 m (2 ft) apart. The experimenter began by reading a short introduction on the inaccuracy of eyewitness testimony. This was followed by one of two manipulations of task importance (see below) and general instructions describing the lineup task and the description task (see above). Participants completed informed-consent forms at this point.

Participants were then briefly questioned by the experimenter to ensure that they correctly comprehended the various instructions. The experimenter then presented the eyewitness identification trials. In the four experimental conditions, the participants saw the same stimuli with the same difficulty manipulations as in the control conditions. Experimental participants, however, called out their responses to each trial orally, and the experimenter recorded the responses of the participant. On all critical trials, the participant answered after the confederates. To dispel suspicion, the order of responding was rotated on the noncritical trials so that the participant answered first on several trials (see Crowne & Liverant, 1963). Following the two eyewitness identification (i.e., conformity) tasks, participants were separated from confederates and asked to complete a brief questionnaire assessing their perceptions of certainty and task importance. They then received a full debriefing and were informed that allocation of the prize money (to 12 students) would be based on a lottery conducted after all sessions were completed.

Manipulation of low task importance. The experimenter read the following instructions to participants in the low-importance condition:

> This is a pilot study whose goal is to try and develop materials to test eyewitness accuracy. It would be useful sometime in the future to develop a test that lets us know how accurate people are as eyewitnesses. Learning more about materials such as the ones you are about to see is a first step in this process. At the moment we need to see how well people can make identifications based on a series of drawings we've developed. We have no useful norms yet on how people respond to the drawings you will see. First we need to explore the best way to present the stimuli. In other words, we are interested in learning more about the optimum conditions for presenting the materials; such as lighting, time of exposure, and distance from the material. Your reactions will give us some useful hints about how to present this material to others.

Manipulation of high task importance. The experimenter read the following instructions to participants in the high-importance condition:

> Because of the importance of accuracy in eyewitness testimony, the University of Iowa is currently upgrading a test we developed several years ago measuring peoples' skill at identifying and remembering details. The Iowa Witness Identification Test, or I.W.I.T., is a test that measures your ability to accurately identify and describe criminal suspects. The I.W.I.T. will soon be adopted by police departments and courtrooms in cases involving eyewitness testimony. Its primary use will be to discern good eyewitnesses from poor ones. Over the next few weeks, we will be running research participants such as yourselves through the revised I.W.I.T. to establish reliable norms. As a result, we hope that you try your best on this test, because establishing accurate norms is crucial. Most people do try hard on this test, because they are interested in seeing

how good their eyewitness accuracy is compared to others. But, to increase your interest in doing well on this test, we will be awarding prizes of $20 at the end of the experimental testing period to the research participants who score the highest in accuracy. Your chances are approximately 1 in 12 of receiving the award.

Dependent measures. Four items were used to assess participant certainty regarding the stimuli in both the baseline control study and in Study 1. These items were included in the questionnaire administered just after the eyewitness tasks and were as follows: "How certain were you of your responses on these tasks?", "How difficult did you find the tasks?", "If you were required to bet between $1 and $7 based on the correctness of your response, what would your average bet be?", and "To what extent were you guessing on these tasks?" Participants responded to these questions after each trial by marking 7-point scales that ranged from *not at all* to *extremely* (with the exception of the betting question).

The final questionnaire also contained three items that measured participants' perceptions regarding task importance. These items were: "How hard did you try on these tasks?", "How much did you care about giving correct responses in this study?", and "How much desire did you have to be above average on these tasks?"

As above, all questions were answered on 7-point scales that ranged from *not at all* to *extremely.* The questionnaires did not ask about each task individually given that we were concerned only with a broad sense of participants' perceptions of difficulty and motivation to be accurate.

The conformity score was defined as the number of trials in which the participant offered the same (incorrect) response as the confederates. The lineup task consisted of seven critical trials (Trials 3, 5, 6, 9, 10, 11, and 13) and six filler trials. A critical trial was one in which the confederates reported a unanimous incorrect choice. As a result, conformity scores in the lineup task could vary from 0 to 7. The description task consisted of seven critical questions and seven filler questions. These were distributed so that for some slides confederates gave two wrong answers, on others one wrong answer, and on others no wrong answers. On the critical questions, the confederates were unanimously wrong. On this task, too, participants could conform (i.e., report the same incorrect answer as the confederates) up to seven times on each task. Total conformity scores for each task were recorded as the mean number of trials (out of a total of seven) in which participants conformed. Conformity scores are also occasionally discussed in terms of the percentage of trials on which conformity occurred (to facilitate comparison to Asch's [1951, 1956] classic research.

Results

Baseline control data. As an initial step, we examined the control study data to check the effectiveness of the difficulty manipulation. For the lineup task in the low-difficulty condition, participants answered correctly on 12.67 out of 13 trials (97%), compared to Asch's (1951) rate of 98% correct. In the high-difficulty condition, accuracy scores dropped to 9.41 out of 13 (72%). In addition, in the low-difficulty condition, participants' mean score on the four certainty items on the baseline control questionnaire (see above) was 5.23 (with 7 indicating *very certain*); the mean rating in the high-difficulty condition dropped to 3.40. We performed analyses of variance (ANOVAs), which revealed that both the percentage correct, $F(1, 28) = 31.07$, $MSE = 2.57$, $p < .001$, and participants' confidence ratings, $F(1, 28) = 29.51$, $MSE = 0.85$, $p < .001$, were significantly lower in the high-difficulty compared with the low-difficulty condition.

On the description task, in the low-difficulty condition, participants answered correctly on 13.53 correct out of 14 trials (97%). In the high-difficulty condition, accuracy dropped to 10.41 out of 14 (74%). An ANOVA indicated that this difference was significant, $F(1, 28) = 75.46$, $MSE = 0.97$, $p < .001$. Nevertheless, given the 72% (lineup task) and 74% (description task) accuracy rates and intermediate confidence scores, it is probably accurate to recognize that our high-difficulty condition was more akin to a moderate level of difficulty.

In addition to this verification of the difficulty manipulation, we examined the control data to measure the baseline tendency to choose those incorrect responses that would later be endorsed by confederates. We refer to these below as *baseline errors.* Measuring these errors permits us to compare the conformity errors induced by confederates in the experimental conditions with these baseline errors. Such baseline errors in the low-difficulty condition were almost nonexistent, with an average of less than 1% on the seven critical trials in both tasks. Baseline errors in the high-difficulty condition also were low, averaging 7% for the lineup task and 26% for the description task.

Lineup task. As an initial step, we made a check on possible gender differences, experimenter effects, and confederate team effects on conformity and confidence. None were significant. Consequently, male and female scores were collapsed within each condition. Conformity rates on the experimental trials for the lineup task are summarized in Figure 2. A 2 (task importance) × 2 (task difficulty) ANOVA revealed a main effect for task difficulty, $F(1, 56) = 10.26$, $MSE = 2.47$, $p < .01$, with more conformity occurring in the high-difficulty condition than in the low-difficulty condition. An interaction between task difficulty and task importance also emerged, $F(1, 56) = 8.26$, $p < .01$. Planned follow-up procedures revealed that on the low-difficulty task, increasing the stakes decreased conformity, $F(1, 56) = 4.37$, $p < .05$, supporting our first hypothesis. In the low-difficulty condition the conformity means were 2.33 (i.e., conformity on 2.33 or 33% of the seven critical trials) for low task importance and 1.13 (16%) for high task importance.

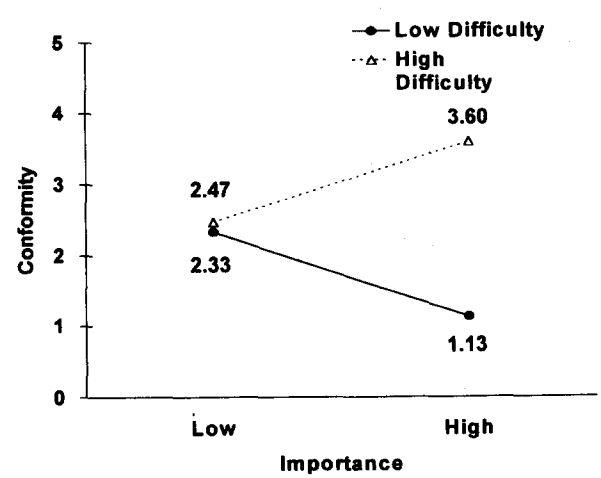

Figure 2. Study 1: Mean number of conforming trials by condition on the lineup task.

These conformity rates (i.e., errors) were significantly greater than control baseline errors in the low-difficulty–control (no confederate) condition (< 1%), $F(1, 84) = 22.53$, $p < .01$, and $F(1, 84) = 4.96$, $p < .05$, for low and high importance, respectively.

In the high-difficulty condition, planned follow-up procedures revealed a significant increase in conformity as task importance increased, $F(1, 56) = 3.88$, $p = .055$, supporting our second hypothesis. The conformity means were 2.47 (35% of trials) for low task importance and 3.60 (51% of trials) for high task importance. These conformity rates were significantly greater than the rate of 7% in the high-difficulty–control (no confederate) condition, $F(1, 84) = 18.90$, $p < .001$, and $F(1, 84) = 45.18$, $p < .0001$, for low and high importance, respectively.

Description task. Conformity rates for the description task are presented in Figure 3. A 2 × 2 ANOVA revealed a main effect for task difficulty, $F(1, 56) = 37.73$, $MSE = 2.42$, $p < .001$, indicating that the difficult condition produced more conformity than the easy condition. An interaction between judgment difficulty and task importance also emerged, $F(1, 56) = 3.97$, $p < .051$. Planned follow-up procedures revealed that in the low-difficult condition, increasing the stakes decreased conformity, $F(1, 56) = 4.46$, $p < .05$, again supporting our first hypothesis. The means were 2.13 (30% of trials) for low task importance and 0.93 (13% of trials) for high task importance. These conformity (error) rates were significantly greater than the baseline rates of less than 1% seen in the low-difficulty (no-confederate) control condition, $F(1, 84) = 17.88$, $p < .001$, and $F(1, 84) = 3.62$, $p < .05$, for low and high importance, respectively. In the high-difficulty condition, there was a tendency for conformity to increase as task importance increased (low importance = 54% of trials; high importance = 60% of trials); however, this effect failed to reach conventional levels of significance, $F(1, 56) = 0.50$, $p = .48$. These conformity (i.e., error) rates for low and high task importance were significantly greater than the baseline error rate of 26% observed in the high-difficulty–control (no-confederate) condition, $F(1, 84) = 14.73$, $p < .001$, and $F(1, 84) = 21.71$, $p < .001$, for low and high task importance, respectively.

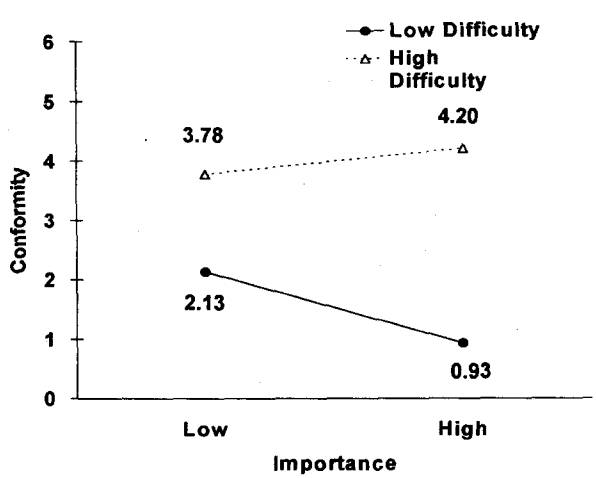

Figure 3. Study 1: Mean number of conforming trials by condition on the description task.

Questionnaire ratings of certainty and importance. At the conclusion of both tasks, the participants were given questionnaires measuring their judgment certainty and the perceived importance of the tasks. On 7-point scales (with higher scores indicating more certainty), participants rated their certainty in the easy (i.e., low-difficulty) condition to be higher than in the difficult condition (Ms = 4.88 and 3.51, respectively), $F(1, 56) = 29.81$, $p < .001$. The participants' ratings of task importance, however, yielded no significant differences between high- and low-important conditions on any of the three measures of task importance or on the sum of these three items, largest $F(1, 56) = 1.30$, $MSE = 9.75$, $p = .30$ (for the summed index). The intercorrelations between these three measures vary from .39 to .41, suggesting a positive but modest relationship. The cell means on this index were as follows: high task importance–low task difficulty, $M = 17.07$; high task importance–high task difficulty, $M = 15.07$; low task importance–low task difficulty, $M = 15.33$; low task importance–high task difficulty, $M = 15.33$.

One possible explanation for these data is that participants may have devalued the importance of the task in the high-importance–high-difficulty condition as a means of rationalizing any poor performance on this more difficult task. In accord with this post hoc interpretation, the first mean does differ marginally from the remaining three means, $F(1, 56) = 3.85$, $p < .06$. Nevertheless, one must concede that the overall data pattern on this index does not provide strong validation for the effectiveness of this manipulation. On the other hand, despite these disappointing manipulation check data, the fact that the importance manipulation did affect conformity rates suggests that the manipulation had some impact on participants. Also, spontaneous comments of participants in postexperimental interviews revealed that in the high-importance conditions (in which financial rewards were mentioned), most participants were excited by the possibility of winning the reward. To explore this issue further, we collected some additional data to more closely examine the impact of the importance manipulation.

Post hoc judgment data regarding task importance. Ninety-four members of an undergraduate social psychology class heard a brief description of Study 1 (described as a study of eyewitness accuracy), saw one set of slides, and learned that the original participants were Introductory Psychology students. These post hoc control participants then were told that the original participants had received instructions that corresponded to the high- (or low-) task-importance instructions. Accordingly, the post hoc controls read either the high- or low-task-importance manipulation and responded to several items assessing task importance. The questions were as follows: "How much effort would the average person (in this study) invest in performing well?", "How motivated would the average person (in this study) be to do his or her best?", "How much would the average person (in this study) desire to be well above average in eyewitness accuracy compared to others in the study (i.e., in the top 12%)?", "How much pressure to perform well would the average person (in this study) experience?", and "How valuable

are the various inducements for performing accurately in this situation?" These five items were rated on 9-point scales that ranged from *not at all* to *extremely.* Three additional questions had a 9-point semantic differential format. These questions were based on the stem "How would the average person (in this study) feel about the eyewitness tasks?" This stem question was followed by three semantic differential scales assessing interest, excitement, and effort (e.g., *not interesting at all–extremely interesting*). This resulted in a total of eight questions. These were summed to form a task importance index. The results indicate that on this index, post hoc participants judged the task to be more important and involving after reading the high-task-importance manipulation ($M = 52.84$, $SD = 6.56$), as opposed to reading the low-task-importance manipulation ($M = 45.74$, $SD = 8.90$), $t(93) = 4.32$, $p < .0001$. This substantially raises our confidence in the effectiveness of the task importance manipulation. This is particularly true given recent reports that judgment studies of this type are quite accurate at estimating manipulated differences in emotion and cognition (e.g., Miller & Carlson, 1990; Miller, Lee, & Carlson, 1991).

Discussion

In Study 1 we examined how task importance affected conformity behavior. We had hypothesized that increasing task importance would have opposite effects on conformity on easy versus difficult tasks. We predicted that on the low-difficulty task, an increase in task importance would lower conformity, whereas on the difficult task increasing task importance would heighten conformity. These hypotheses were generally supported by the data (although the positive relationship between conformity and task importance under high task difficulty was found only on the lineup task). It is noteworthy that the conformity rates (33% of trials for the lineup task and 30% of trials for the description task) in the easy condition (without any explicit incentives or rewards) closely approximate the rates originally reported by Asch (1956; 33% of trials) with a judgmental task of roughly this difficulty. When incentives for accuracy were present, this rate dropped ($p < .05$) to 16% and 13% in the lineup and description tasks, respectively. Note, however, that this still represents a significant ($p < .05$) degree of social influence when compared with the error rate (< 1% in both tasks) of controls. This speaks to the power of social influence. In this condition, conformity occurred in 16% of lineup trials and 13% of description task trials, despite the fact that the group norm was incorrect, the correct answer was obvious, and that there were meaningful incentives for accuracy. In short, under conditions of extremely low difficulty, increasing task importance by offering psychological and financial incentives for accuracy significantly lowered, but did not completely eliminate, social influence. Our confidence in these data is enhanced by the fact that Di Vesta (1959) reported a similar effect in a study in which a task of low difficulty was used.

The moderating effect of task difficulty is reflected by the fact that the data pattern just described is reversed under conditions of moderate task difficulty. Specifically, when task difficulty was increased, high task importance increased the conformity observed ($p < .06$), at least on the lineup task. As noted, our interpretation of these data is that, when it is important to be accurate and the task at hand is difficult, participants turn to social feedback from others in an attempt to obtain cues regarding the most apparently accurate response. As noted, this is closely akin to Festinger's (1954) classic view regarding the important role of social reality in cases involving low objective reality.

It is our feeling that this high-difficulty *incentive effect* is the more interesting of the two simple effects comprising the Importance × Difficulty interaction observed on these conformity tasks. That is, it is not particularly surprising that substantial incentives for accuracy lead people to resist influence from a group of obviously inaccurate confederates, as was observed in low-task-difficulty conditions. If people are paid for accuracy, they become more accurate! As a result, the opposite outcome in the high-difficulty conditions is both counterintuitive, and personally counterproductive, for participants. The more important it is for them to be accurate, the more they are influenced by an inaccurate group consensus. As noted, we think this outcome has important empirical as well as pragmatic implications.

Therefore, one disquieting note is that task importance (i.e., incentive) produced an increase in conformity in the difficult condition only on the lineup task and then only at a borderline level of statistical significance. As a result, we conducted Study 2 to permit replication and an examination of additional variables. Given the robust (and replicated) results observed in the easy (low-difficult) condition (Di Vesta, 1959), and the desire to increase power, Study 2 focused on only difficult conditions. In this study we also eliminated the description task and added several critical and filler trials to the lineup task in an attempt to increase its reliability and power.

Moreover, Study 2 addressed an additional issue: the relationship between conformity and postjudgment confidence. The analysis of social influence processes, offered above, suggests that when judgment difficulty is high, social cues become increasingly important as indicators of judgmental correctness (Festinger, 1954). If so, on ambiguous judgments individuals may actually feel greater postresponse confidence in the accuracy of their judgments having conformed to an *inaccurate group norm* than those who fail to conform. The thinking here is that conformers have the comfort of social support, whereas nonconformers have the scant comfort of their own unsupported perception of a degraded or ambiguous stimulus. This effect, however, should be most apparent when the task judgment is difficult and the norm is strong, that is, when other group numbers are unanimous and confident about their opinions. To examine this issue, in Study 2 we manipulated the confidence expressed by confederates and measured the postjudgment confidence of participants after each trial. Both the issue and the manipulation are relatively unique. A good number of studies have examined how prejudgment confidence is related to eventual conformity. This research follows the strategy of manipulating either task difficulty (e.g., Deutsch & Gerard, 1955) or participant confidence on a prior task (Julian, Regula, & Hollander, 1968) and generally finds that when initial confi-

dence can be inferred to be high (e.g., in easy conditions), subsequent conformity is lower.

Several other studies have examined how conformity relates to an overall retrospective confidence score (collapsed over critical and noncritical trials) obtained after all conformity trials are over (Campbell, Tesser, & Fairey, 1986; Hancock & Sorrentino, 1980; Kelley & Lamb, 1957; Meunier & Rule, 1967; Penner & Davis, 1969; Rule & Sandilands, 1969; Savell & Healey, 1969; Tesser, Campbell, & Mickler, 1983). The general finding in this research is that conformity and confidence are negatively related. This may be due to the fact that this research may have focused on judgments of low difficulty (e.g., Campbell et al., 1986; Hancock & Sorrentino, 1980).[4] An alternative explanation is that this research, with the exception of a study by Crowne and Liverant (1963), used single-item retrospective measures of confidence. Crowne and Liverant (1963) did obtain confidence scores after each trial in one of their three conditions and found that, here too, confidence and conformity were negatively related. As noted elsewhere (see footnote 2), however, there is some question regarding the difficulty level of the judgment task used in this study. A second problem is that, in this study, the betting data (obtained in another condition) did not parallel the confidence reports. Highly compliant participants (based on a median split) actually bet more (thereby showing greater confidence) on trials on which they conformed than on trials on which they did not conform. In sum, it is hard to draw confident conclusions on the basis of these data. With the exception of Crowne and Liverant's (1963) study, we are unaware of other studies in which confidence was measured after each trial (particularly on judgments of high difficulty), and we are not aware of any studies that manipulated confederate confidence. In Study 2 we addressed both issues.

Study 2

Method

Overview. Study 2 was a 2 (task importance) × 2 (confederate confidence) between-subject factorial with one additional baseline error control cell. As in Study 1, this control condition allowed us to assess the frequency with which individuals who responded privately happened to make precisely the same errors made by the confederates in the experimental conditions. This then established a baseline error (or conformity) score. All participants completed 20 trials of the lineup task (under high-difficulty conditions) from Study 1 as well as confidence ratings of each judgment.

Participants. Participants were 99 Introductory Psychology students (45 men, 54 women) who participated in the study as one means of fulfilling a course requirement. Four additional individuals were excluded because they surmised the general nature of the experiment (i.e., that it involved a nonindependence of judgment). Twenty-one students were randomly assigned to the baseline control cells. Control participants were run in groups of 3 but, unlike experimental participants, they made their judgments privately in writing. The remaining 78 individuals were randomly assigned to experimental conditions.

Materials. The stimuli and instructions were the same as those used in the lineup task from Study 1. Seven additional sets of perpetrator/lineup drawings were added, resulting in 40 slides in all. This permitted 10 filler trials and 10 critical trials (Trials 4, 6, 8, 9, 10, 12, 15, 17, 19, and 20).

Level of judgment difficulty. All participants saw the slides under the same conditions used in the high-difficulty condition in Study 1, that is, a single 0.5-s exposure per slide.

Manipulation of task importance. This manipulation (and the general cover story for the experiment) was the same as in Study 1.

Procedure. The procedure in Study 2 was the same as that used in Study 1, with the following exceptions. First, participants privately indicated their (postjudgment) confidence in each of their (oral) judgments on a scale that ranged from 1 (*not confident at all*) to 7 (*completely confident*). These written responses were made just after each trial. The internal consistency reliability (coefficient alpha) for these ratings was .8. Second, to further camouflage the purpose of the experiment, a tape recorder playing a jazz selection (Miles Davis: "Kind of Blue") was turned on by the experimenter (at a moderately low level) in all conditions as the lineup task began. This ostensible manipulation was intended to divert attention from the study's true purpose by subtly suggesting to participants that the study concerned the effect of music on their eyewitness performance. Third, in one set of conditions (high confederate confidence), confederates projected high confidence when offering their estimates across filler and critical trials, whereas in the other (low-confidence) condition, confederates projected a lack of confidence on both critical and filler trials (see next section). A fourth difference in the procedure of Study 2 is that participants were not asked their perceptions of task importance, given the insensitivity of these items in Study 1. A fifth difference was that the description task was dropped from the procedure. This was based both on its insensitivity in Study 1 (in the high-difficulty condition) as well as on the fact that it differed qualitatively from standard conformity measures that require either a continuous judgment (regarding number, distance, movement, etc., e.g., Sherif, 1935) or matching a stimulus to a standard (e.g., Asch, 1956). In this regard, the lineup task seemed more directly analogous to the classic Asch-type task. After participating in the 20 lineup task trials, participants were thanked and debriefed. At the end of the Study 2, participants were selected by lottery to receive a $20 payment.

Manipulation of confederate confidence. In the low-confederate-confidence condition, confederates often hesitated before offering answers, used a tentative tone of voice, often offered their answers as questions (e.g., "Number 3?"), used verbal qualifiers (e.g., "Well . . . number 3, maybe"), verbally expressed uncertainty (e.g., "Gee I don't know . . . number 3, I guess"), used filled pauses ("Aaah, umm") and a variety of other nonverbal signs of uncertainty and confusion such as vigorously exhaling on seeing the slide ("Phew"), head shaking, leaning forward and peering at the slide, and so on. These low-confidence behaviors were enacted on an ad lib basis by confederates so that confederates were not repeating themselves precisely from trial to trial. On each low-confidence trial (both critical and noncritical), each confederate was instructed to express a lack of confidence through the use of two to three low-confidence cues from the array described just above. In the high-confederate-confidence condition, confederates offered their judgments with little delay, in a confident tone of voice, using no qualifying remarks or filled pauses. These differences in confederate behavior took place on all trials. In practice sessions there was 100% agreement among six observers regarding when the confederates were implementing the low- versus the high-confidence manipulation over 10 trials.

[4] With few exceptions (e.g., Hancock & Sorrentino, 1980), generally these published reports offer few details regarding task difficulty, particularly regarding baseline error rates. As a result, it is hard to be certain regarding the difficulty used in this research. In some cases, however, the nature of task exposure (e.g., the use of unspeeded judgments; Campbell et al., 1986) suggests that the task judgments were often not overly difficult.

Results

Preliminary analyses. As in Study 1, an initial analysis assessed possible differences involving participant gender and experimenter effects. We did not analyze confederate gender, given that the great majority of sessions involved mixed gender confederate teams (n = 61) with only four sessions run with all-male teams and 13 with all-female teams. In Study 1, however, this factor did not produce significant differences in either conformity or confidence. The preliminary analyses showed no significant main effects or interactions for experimenter or experimenter gender. Participant gender had no significant effect on overall rates of conformity in experimental conditions (male conformity M = 3.48; female conformity M = 3.65; $F < 1$, ns). Similarly, there were no differences between men and women in the baseline control condition, with both genders having almost identical rates of baseline "conformity" errors and rates of confidence (men's "conformity" errors = 1.08, women's = 1.00; confidence: male M = 4.74, female M = 4.73; $F < 1$). In the four experimental conditions, however, there was a general tendency for men to report more postjudgment confidence on both critical and noncritical trials (Ms = 4.85 and 5.20, respectively) than women (Ms = 4.25 and 4.41, respectively), critical trials $F(1, 55) = 3.56$, $p = .067$, neutral trials $F(1, 55) = 5.92$, $p = .019$. Most critically, however, men and women were distributed approximately equally across all four experimental conditions, and participant gender did not come close to interacting with treatments on either confidence or conformity (all Fs < 1). Given this data pattern, the results reported below collapse over participant gender, experimenter, and confederate gender.

Conformity. We analyzed the conformity and confidence data with a 2 (task importance) × 2 (confederate confidence) × 2 (trial block, i.e., first five trials vs. second five) ANOVA with the first two factors being between-subject effects and the last a within-subject effect. This analysis yielded three significant main effects. Conformity was greater when confederates expressed confidence (M = 4.54) than when they did not (M = 2.95), $F(1, 75) = 7.78$, $MSE = 6.10$, $p = .004$, one-tailed. Conformity was greater when there were incentives for task accuracy (M = 4.30) than when there were no such incentives (M = 3.24), task importance $F(1, 75) = 3.58$, $p < .031$, one-tailed, and conformity was greater in the first block of five trials (M = 2.04) than in the second block of five trials (M = 1.70), trial block $F(1, 75) = 6.09$, $p < .02$, two-tailed. No other main effects or interactions were significant. See Table 1 for conformity means by condition.

In a supplementary analysis we compared the four experimental means and one control mean in an ANOVA with condition (5 levels) and trial blocks (2 levels) as factors. The condition effect was highly significant, $F(4, 89) = 8.95$, $p < .0001$. Follow-up tests with Dunnette's F statistic (for testing all conditions against a control condition, cf. Winer, 1962) indicated that the mean conformity errors in the low-importance–low-confederate-confidence condition did not differ from the baseline error level observed in the control condition, $t(89) = 1.70$. The mean conformity errors in all three remaining experimental conditions, however, was significantly greater ($p < .01$ in all cases) than the baseline "conformity" errors observed in the control condition, lowest $t(89) = 3.35$, Dunnette's critical values for $p > .05 = 2.55$, and for $p < .01 = 3.04$. There was also a significant effect of trial blocks, $F(1, 89) = 8.18$, $p < .005$, reflecting the fact that conformity was greater on the first block of trials (see above).

Table 1
Study 2: Conformity Means by Condition

Task importance	Confederate confidence: Low — Score	Low — n	Low — SD	High — Score	High — n	High — SD
Low	2.33	18	1.33	4.05	20	2.37
High	3.53	19	2.80	5.00	21	3.36

Note. Baseline error in control condition = 1.04 (n = 21, SD = 1.02). Maximum possible conformity score = 10. As a result, the percentage of conforming trials is computed by moving the decimal point one place to the right, that is, conformity occurred on 50% of trials in the high–high cell. Baseline error rate reflects the frequency with which baseline control participants made the same errors as those endorsed by confederates in the experimental cells.

Confidence on critical trials and confidence on neutral trials. As noted, we analyzed the confidence data in the four experimental conditions in the same manner as the conformity data (i.e., a 3-way ANOVA). There were no main effects or interactions that approached significance for either confidence on critical trials or confidence on neutral trials (all Fs < 1). The supplementary ANOVA, which included the baseline-control cell with the four experimental cells, revealed a significant main effect of trial block, $F(1, 89) = 6.24$, $p = .014$, on confidence on neutral trials. This main effect was qualified by a (5) Condition × (2) Trial Blocks interaction, $F(4, 89) = 2.92$, $p = .025$. These effects reflect the fact that, on neutral trials, confidence in the four experimental conditions (M = 4.72) generally was higher than confidence in the control condition (M = 4.10) and that, because of an increase in confidence in the four experimental conditions in Trial Block 2, this effect was significantly more pronounced in the second block of trials (control M = 4.05 vs. average of experimental means M = 4.89). Thus, apparently the availability of confederates as sources of social comparison raised the confidence of experimental participants on neutral trials (on which confederate responses were accurate). As noted, this same pattern was apparent on Trial Block 1 (ns) and over all trials, $F(4, 89) = 2.043$, $p < .10$ (see Table 2).

Correlations between conformity and confidence. One must remember in interpreting the null confidence data on critical trials that our interest in confidence is primarily in its relationship to conformity. Although we would have been encouraged if confidence on critical trials was higher in the experimental cells (especially those involving confident confederates) than in the control cell, this prediction was undermined by the fact that on at least 50% of trials conformity did not occur. As a result, participants disagreed with others on such trials. This type of disagreement is unlikely to heighten confidence. Therefore, it is essential to examine the correlations between conformity and

Table 2
Confidence on Neutral Trials by Condition

Task importance	Confederate confidence: Low Score	Low *n*	Low *SD*	High Score	High *n*	High *SD*
Low	4.51	18	.98	4.94	20	1.00
High	4.68	19	1.14	4.72	21	1.40

Note. Mean confidence in baseline control condition was 4.10 ($n = 21$, $SD = .76$). High scores = high confidence (possible range: 1–7).

confidence. These correlations are presented by condition in Table 3.

As can be seen in Table 3, when confederates expressed confidence in their judgments (i.e., in the high-confidence condition), there was a moderately strong positive correlation between participants' degree of conformity and their postjudgment confidence on critical trials (mean $r = .43, p < .005$). This positive relationship was reversed when confederates expressed a lack of confidence. The average correlation (using the r-to-Z table) in the two low-confidence cells is $-.33$ ($p < .025$). Correlations between conformity and confidence on neutral trials tended to produce a similar data pattern, albeit somewhat less significant and robust (average r in low-confidence conditions = $-.20, p < .10$; average r in high-confidence condition = $.47, p < .05$). The positive correlation in the high-confederate-confidence conditions indicates that participants reported greater conformity in these conditions the more they conformed to an inaccurate group consensus.

Effect size and joint probability data. Given the convergent procedures used in Study 1 and Study 2 to manipulate task importance and to assess its affects on conformity, meta-analytic statistics regarding estimates of effect size and combined probability values are presented in Table 4. These data refer to outcomes on the lineup task under high-difficulty conditions.

As Table 4 indicates, task importance produced a modest yet quite statistically reliable impact on conformity across Studies 1 and 2—at least, as measured by the lineup task. The mean effect size of .23 is generally viewed as a small effect (J. Cohen, 1977), yet the combined p value ($p = .004$) is quite respectable.

Table 3
Study 2: Correlations Between Postjudgment Confidence on Critical Trials and Conformity by Condition

Task importance	Confederate confidence: Low *r*	Low *p*	High *r*	High *p*
Low	−.035	.889	.424	.002
High	−.573	.01	.437	.041

Note. Positive correlations indicate that greater conformity is associated with greater postjudgment confidence. The p values indicate whether correlations differ from 0.

Table 4
Meta-Analytic Data for Studies 1 and 2

Study	*df*	*r*	*Zr*	*p*	*Z*
1	56	.255	.261	.028	1.92
2	75	.213	.216	.031	1.87

Note. Only the data for the difficult conditions in Study 1 are included. Rosenthal (1991) recommended the use of one-tailed tests in cases where outcomes are consistently in a predicted direction, as in the present research. Combined effect size estimates: weighted combined Zr (by df) = .235, $r = .23$; unweighted combined $Zr = .239$, $r = .235$. Combined Z and p values: weighted Z (by df) = 2.647, $p = .004$, one-tailed; unweighted $Z = 2.68$, $p = .004$, one-tailed. $Zr = r$ to Z transformation.

General Discussion

The data of Study 2 add substantially to our confidence regarding the impact of task importance on conformity under conditions of moderate to high judgment difficulty. Study 2 replicated the counterintuitive finding of Study 1 that heightening incentives for accuracy actually heightened participants' susceptibility to an inaccurate group consensus. Indeed, in Study 2 this effect was not significantly attenuated when confederates indicated a general lack of confidence in their own judgments—a condition that we originally thought might moderate the phenomenon. Our explanation for these effects echoes Festinger's (1954) arguments that when judgments cannot be verified objectively, individuals become increasingly reliant on social information to gauge the accuracy and appropriateness of their views. As a result, if the motivation for accuracy increases on such ambiguous judgments (whether due to incentives or threats), one should be more likely to rely on such social cues given that they often will be the only cues available regarding accuracy (other than the inadequate stimulus cues). Such reliance should be particularly strong when social cues seem most reliable (i.e., when others are in consensus and seem confident and capable regarding their judgments). The fact that in Study 2 even nonconfident confederates had greater impact under conditions of high task importance may be due to the fact that such confederates were in agreement. As a result, this consensus may have been sufficiently convincing to outweigh their individual expressions of low confidence.[5]

In conditions in which confederates expressed low confidence, one might be tempted to explain the relatively higher conformity of participants in the high- (vs. low-) importance condition in terms of an ineffective confidence manipulation. Given the large overall effect of confederate confidence on con-

[5] As noted in footnote 2, one could also easily rephrase this explanation using modern theories of persuasion. Thus, Chaiken (1987) and Petty and Cacioppo (1986) could characterize high task importance as a manipulation of issue involvement that increases the desire for central (or systematic) processing. This would lead to heightened reliance on social cues given the unavailability (and unreliability) of objective information about the stimuli (Maheswaren & Chaiken, 1991; cf. Eagly & Chaiken, 1993, for a discussion of accuracy motivation as a variable in persuasion).

formity, however ($F = 7.78$), coupled with the broad and salient confederate behaviors involved in the manipulation, we think that this explanation is implausible. Note that the strong impact of confederate confidence on conformity in Study 2, though not unexpected, is a unique finding in the conformity/social influence literature. We have been able to find no other research to date investigating how this variable affects conformity.

One alternative explanation to the social influence explanation we outline just above is that our results represent some form of a choking effect (e.g., Baumeister, 1984). That is, with the difficult task data, it is conceivable that the greater conformity observed in Studies 1 and 2 under high task importance simply reflect the fact that the task importance manipulation increased task pressure on participants, thereby elevating errors in general and making participants more responsive to erroneous social influence from confederates in particular.[6] Fortunately, there is a means of testing this alternative view. If choking effects are occurring *as opposed to* social influence effects, incentive should elevate a variety of errors in addition to the errors advocated by the confederates. In addition, this should be true on noncritical trials as well as on critical trials. A post hoc analysis of errors other than those advocated by the confederates (i.e., independent errors) indicated no support for this view. In these analyses, *independent errors* are defined as the sum of those errors on critical trials that do not agree with confederate responses and any errors occurring on noncritical trials (in which confederates offer correct responses). When these independent errors serve as the dependent variable, there are no choking effects (i.e., more errors under greater incentive) that approach significance on either the lineup task (in Studies 1 and 2) or the description task (in Study 1).[7] Indeed, contrary to the choking hypothesis, the largest difference found in these analyses (on the lineup task in Study 1) indicates that independent errors of this type are more likely to occur in the low-task-importance condition ($M = 2.53$ of a possible 13) rather than in the high-task-importance condition ($M = 1.67$), $F(1, 28) = 2.56$, $p = .13$. In short, the presence of both ego-based and financial incentives (in the high-importance conditions) did not increase the presence of independent errors as one would expect if errors in Studies 1 and 2 represented a form of choking. As a result, we think the data of Studies 1 and 2 are best viewed as social influence phenomena. This interpretation is supported by the correlational relationship between conformity and participant confidence reported in Study 2.

The positive correlations observed in the high-confederate-confidence conditions between participants' postjudgment confidence and whether participants conformed on a particular trial are congruent with Festinger's (1954) notion that conformity (i.e., achieving group unanimity) often serves as a means of reducing uncertainty. Of course, it is hard to establish causality from these correlational data, but it is interesting that when perceived group consensus is experimentally manipulated, its presence serves to heighten participant confidence (Baron et al., in press; Goethals & Nelson, 1973; Orive, 1988). In the present research, however, this uncertainty reduction effect was either eliminated or significantly reversed in cases where confederates appeared to have doubts about their own judgments. Thus, in the high-importance–low-confederate-confidence condition, conformity was associated with lower confidence ($r = -.57$). Given that participants in this condition responded with relatively greater conformity (35%) than that exhibited by participants in the low-importance–low-confederate-confidence conditions (23%), it is hard to characterize the impact of task importance on conformity solely as a simple uncertainty reduction process. In these cells, conformity tended to increase as a function of task importance despite the fact that such conformity (to a hesitant and incorrect group of confederates) led to an increase in participant uncertainty. Here it seems that, rather than the individual being confident that the group is correct, participants are tentatively hoping (despite their doubts) that "all those people can't be wrong." The fact that their conformity tends to increase as task accuracy becomes more important despite these misgivings suggests that the cues provided by social consensus represent some of the only hints participants can access regarding judgment accuracy. Thus, these cues must be used if accuracy is crucial.

On the other hand, the positive conformity–confidence correlations observed in conditions where confederates acted with confidence and enthusiasm indicates that not only will a majority consensus often be capable of misleading individuals into inaccurate, irrational, or unjustified judgments, but that such consensus can produce heightened confidence in such judgments as well. These data then provide insight into social influence techniques seen in a variety of settings, from TV's gushy group infomercials to religious indoctrination where care is taken to surround recruits with nothing but enthusiastic "guru-struck" disciples. Our data suggest that so long as the judgments are difficult or ambiguous, and the influencing agents are united and confident, increasing the importance of accuracy will heighten confidence as well as conformity—a dangerous combination.

Although our discussion has emphasized how increasing task importance (motivation for accuracy) can lead to increased conformity, the results from Study 1 make it clear that such an effect will be distinctly moderated by task difficulty. In that study, in two separate tasks, high task importance (i.e., incentives for accuracy) led to less conformity provided that task difficulty was low. At one level, this is a sensible and straightforward prediction in that people are more likely to resist an obviously incorrect group norm when there are several meaningful (ego-related and financial) incentives to be accurate. Nevertheless, those of us who wish to defend the power (and importance) of social influence can point to the fact that on both the lineup and description tasks, the number of confederate-induced errors, though low in these conditions, still significantly exceeded the base rate for such errors established in the control cells. Thus, even on fairly easy and obvious task judgments, incentives for accuracy do not completely eliminate the impact of

[6] We are grateful to an anonymous reviewer for suggesting this line of inquiry.

[7] Similar nonsignificant results occur when independent errors in critical and noncritical trials are analyzed separately. The highest F in these analyses occurred in Study 2 in the analysis of independent errors occurring on critical trials (high-task-importance errors $M = 2.56$; low-task-importance errors $M = 2.36$), $F(1, 74) = 1.24$, $p = .30$.

social influence even when that influence is misguided and inaccurate. One needs an even stronger cautionary note in discussing the implications of the results found with tasks of high judgmental difficulty. These data belie any optimistic view that illogical, inaccurate, emotional, or nonfactual social influence attempts will fall flat when such a judgment has important consequences. The present results suggest that given even a modest degree of ambiguity, increasing the importance of a given judgment offers no corrective influence against an inaccurate majority and, if anything, appears to increase their influence. From one perspective, this is good news in that it documents that social influence phenomena are powerful effects with intriguing counterintuitive pragmatic implications. From a more basic perspective, however, it is disquieting because it illustrates that, given a modicum of ambiguity, our susceptibility to social manipulation is not minimized just because we have substantial vested interest in an issue or judgment. Rather, that vested interest appears to be capable of enhancing the impact of such influence both in terms of conformity and subjective confidence concerning that conformity.

References

Asch, S. E. (1951). Effects of group pressure upon the modification distortion of judgments. In H. Guetzkow (Ed.), *Groups, leadership and men* (pp. 177–190). Pittsburgh, PA. Carnegie Press.

Asch, S. E. (1956). Studies of independence and submission to group pressure: I. A minority of one against a unanimous majority. *Psychological Monographs, 70* (9, Whole No. 417).

Back, K. W. (1951). Influence through social communication. *Journal of Abnormal and Social Psychology, 46,* 9–23.

Baron, R. S. (1986). Distraction conflict theory: Progress and problems. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 19, pp. 1–40). New York: Academic Press.

Baron, R. S., Hoppe, S., Kao, C. F., Brunsman, B., Linneweh, B., & Rogers, D. (in press). Social corroboration and opinion extremity. *Journal of Experimental Social Psychology.*

Baron, R. S., Kerr, N., & Miller, N. (1992). *Group process, group decision, group action.* London: Open University Press.

Baumeister, R. F. (1984). Choking under pressure: Self-consciousness and paradoxical effects on skilled performance. *Journal of Personality and Social Psychology, 46,* 610–620.

Berkowitz, L., & Macaulay, J. R. (1961). Some effects of differences in status level and status stability. *Human Relations, 14,* 135–147.

Campbell, J. D., Tesser, A., & Fairey, P. J. (1986). Conformity and attention to the stimulus: Some temporal and contextual dynamics. *Journal of Personality and Social Psychology, 51,* 315–324.

Chaiken, S. (1987). The heuristic model of persuasion. In M. P. Zanna, J. M. Olsen, & C. P. Herman (Eds.), *Social influence: The Ontario symposium* (Vol. 5, pp. 3–39). Hillsdale, NJ: Erlbaum.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (rev. ed.). New York: Academic Press.

Cohen, S. (1978). Environmental load and the allocation of attention. In A. Baum, J. E. Singer, & S. Valins (Eds.), *Advances in environmental psychology. Vol. 1: The urban environment* (pp. 1–29). Hillsdale, NJ: Erlbaum.

Crowne, D. P., & Liverant, S. (1963). Conformity under varying conditions of personal commitment. *Journal of Abnormal and Social Psychology, 66,* 547–555.

Darley, J. M. (1966). Fear and social comparison as determinants of conformity behavior. *Journal of Personality and Social Psychology, 4,* 73–78.

Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *Journal of Abnormal and Social Psychology, 51,* 629–636.

Di Vesta, F. J. (1959). Effects of confidence and motivation on susceptibility to informational social influence. *Journal of Abnormal and Social Psychology, 59,* 204–209.

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes.* Orlando, FL: Harcourt Brace Jovanovich.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 5,* 327–346.

Goethals, G. R., & Nelson, R. E. (1973). Similarity in the influence process: The belief–value distinction. *Journal of Personality and Social Psychology, 25,* 117–122.

Hancock, R. D., & Sorrentino, R. M. (1980). The effects of expected future interaction and prior group support on the conformity process. *Journal of Experimental Social Psychology, 16,* 261–269.

Hoffer, E. (1951). *The true believer.* New York: Harper & Row.

Janis, J. L. (1972). *Victims of groupthink.* Boston: Houghton-Mifflin.

Jones, E. E., & Gerard, H. B. (1967). *Foundations of social psychology.* New York: Wiley.

Julian, J. W., Regula, C. R., & Hollander, E. P. (1968). Effects of prior agreement by others on task confidence and conformity. *Journal of Personality and Social Psychology, 9,* 171–178.

Kelley, H. H., & Lamb, T. W. (1957). Certainty of judgment and resistance to social influence. *Journal of Abnormal and Social Psychology, 55,* 137–139.

Krech, D., Crutchfield, R. S., & Ballachey, E. L. (1962). *Individual in society.* New York: McGraw-Hill.

Kruglanski, A. W. (1989). *Lay epistemics and human knowledge.* New York: Plenum.

Le Bon, G. (1960). *The crowd.* New York: Viking Press. (Original work published 1895)

Levine, J. M. (1989). Reaction to opinion deviance in small groups. In P. Paulus (Ed.), *Psychology of group influence: New perspectives* (pp. 375–430). Hillsdale, NJ: Erlbaum.

Maheswaren, D., & Chaiken, S. (1991). Promoting systematic processing in low motivation settings. *Journal of Personality and Social Psychology, 61,* 13–25.

Malof, M., & Lott, A. J. (1962). Ethnocentricism and the acceptance of Negro support in a group pressure situation. *Journal of Abnormal and Social Psychology, 65,* 254–258.

Meunier, C., & Rule, B. G. (1967). Anxiety, confidence, and conformity. *Journal of Personality, 35,* 498–504.

Miller, N., & Carlson, M. (1990). Valid theory testing meta-analyses further question the Negative State Relief model of helping. *Psychological Bulletin, 107,* 215–225.

Miller, N., Lee, J. Y., & Carlson, M. (1991). The validity of inferential judgements when used in theory testing meta-analysis. *Personality and Social Psychological Bulletin, 17,* 335–343.

Nicholson, N., Cole, S. G., & Rocklin, T. (1985). Conformity in the Asch situation: A comparison between contemporary British and U.S. university students. *British Journal of Social Psychology, 24,* 59–63.

Orive, R. (1988). Group consensus, action immediacy, and opinion confidence. *Personality and Social Psychology Bulletin, 14,* 573–577.

Penner, L. A., & Davis, J. H. (1969). Conformity and the "rational" use of unanimous majorities. *Journal of Social Psychology, 78,* 299–300.

Perrin, S., & Spencer, C. P. (1981). Independence or conformity in the Asch experiment as a reflection of cultural and situational factors. *British Journal of Social Psychology, 20,* 205–209.

Petty, R., & Cacioppo, J. T. (1986). The elaboration likelihood model

of persuasion. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 19, pp. 123–205). New York: Academic Press.

Rosenthal, R. (1991). *Meta-analytical procedures for social research* (rev. ed). Newbury Park, CA: Sage.

Rule, B. G., & Sandilands, M. L. (1969). Test anxiety, confidence, commitment, and conformity. *Journal of Personality and Social Psychology, 37,* 460–467.

Savell, J. M., & Healey, G. W. (1969). Private and public conformity after being agreed and disagreed with. *Sociometry, 32,* 315–329.

Sherif, M. (1935). A study of some social factors in perception. *Archives of Psychology, 27* (No. 187).

Shirer, W. (1960). *The rise and fall of the Third Reich.* New York: Fawcett.

Tesser, A., Campbell, J., & Mickler, S. (1983). The role of social pressure, attention to the stimulus, and self-doubt in conformity. *European Journal of Social Psychology, 13,* 217–233.

Weightsman, J. M. (1984). *Making sense of the Jonestown suicides: A sociological history of the People's Temple.* New York: E. Mellon Press.

Wilder, D. A. (1977). Perception of groups, size of opposition, and social influence. *Journal of Experimental Social Psychology, 13,* 253–268.

Winer, B. J. (1962). *Statistical principles and experimental design.* New York: McGraw-Hill.

Received June 28, 1995
Revision received May 7, 1996
Accepted May 9, 1996 ■

AMERICAN PSYCHOLOGICAL ASSOCIATION
SUBSCRIPTION CLAIMS INFORMATION

Today's Date:_____

We provide this form to assist members, institutions, and nonmember individuals with any subscription problems. With the appropriate information we can begin a resolution. If you use the services of an agent, please do **NOT** duplicate claims through them and directly to us. **PLEASE PRINT CLEARLY AND IN INK IF POSSIBLE.**

PRINT FULL NAME OR KEY NAME OF INSTITUTION

MEMBER OR CUSTOMER NUMBER (MAY BE FOUND ON ANY PAST ISSUE LABEL)

ADDRESS

DATE YOUR ORDER WAS MAILED (OR PHONED)

____PREPAID ____CHECK ____CHARGE
CHECK/CARD CLEARED DATE:_____

CITY STATE/COUNTRY ZIP

(If possible, send a copy, front and back, of your cancelled check to help us in our research of your claim.)

YOUR NAME AND PHONE NUMBER

ISSUES: ____MISSING ____DAMAGED

TITLE	VOLUME OR YEAR	NUMBER OR MONTH

Thank you. Once a claim is received and resolved, delivery of replacement issues routinely takes 4–6 weeks.

(TO BE FILLED OUT BY APA STAFF)

DATE RECEIVED: _____ **DATE OF ACTION:** _____
ACTION TAKEN: _____ **INV. NO. & DATE:** _____
STAFF NAME: _____ **LABEL NO. & DATE:** _____

Send this form to APA Subscription Claims, 750 First Street, NE, Washington, DC 20002-4242

PLEASE DO NOT REMOVE. A PHOTOCOPY MAY BE USED.