

# Reacting to an Assumed Situation vs. Conforming to an Assumed Reaction: The Role of Perceived Speaker Attitude in Vicarious Dissonance

**Benoît Monin**  
*Stanford University*

**Michael I. Norton**  
*Massachusetts Institute of Technology*

**Joel Cooper**  
*Princeton University*

**Michael A. Hogg**  
*University of Queensland*

Witnessing an ingroup member acting against his or her belief can lead individuals who identify with that group to change their own attitude in the direction of that counterattitudinal behavior. Two studies demonstrate this vicarious dissonance effect among high ingroup identifiers and show that this attitude change is not attributable to conformity to a perceived change in speaker attitude. Study 1 shows that the effect occurs—indeed, is stronger—even when it is clear that the speaker disagrees with the position espoused, and Study 2 shows that foreseeable aversive consequences bring about attitude change in the observer without any parallel impact on the perceived attitude of the speaker. Furthermore, the assumption that vicarious dissonance is at heart a group phenomenon is supported by the results indicating that attitude change is not impacted either by individual differences in dispositional empathy or measures of interpersonal affinity.

**KEYWORDS** cognitive dissonance, conformity, vicarious processes

WHAT happens when individuals witness an ingroup member bow to situational pressures? Imagine witnessing a department colleague deriding the experimental method in an effort to woo a group of cultural anthropologists.

---

#### **Author's note**

Address correspondence to Benoît Monin, Department of Psychology, Jordan Hall, Stanford University, Stanford, CA 94305, USA [email: monin@stanford.edu]

Individuals do take into account situations when attributing causes for behavior (Kelley, 1967), but decades of research on the fundamental attribution error (Gilbert & Malone, 1995; Ross, 1977) have shown that observers rarely feel that knowledge of situational factors is enough to explain an actor's behavior satisfactorily: you may understand why your colleague is taking this position, but still wonder why she needed to sell out *to such an extent*. To observers, responses to situational forces often seem out of proportion with the amount of pressure, because it is difficult for observers to gauge the power of the situation given their removal from it (Nisbett & Ross, 1980). These attributional biases can sometimes have a direct impact on the observer's own views. When the person involved is a member of an important ingroup, the uneasiness inspired by the scene may be such that you end up changing your *own* attitude to match the contrived behavior of the actor (Norton, Monin, Cooper, & Hogg, 2003). If your department is an important part of your identity, upon seeing your colleague placating anthropologists, you may find it hard not to picture yourself in the same situation, imagining how uncomfortable you would feel adopting the same stance; as suggested by decades of research on cognitive dissonance (Festinger, 1957; Harmon-Jones & Mills, 1999), one way to cope with this psychological discomfort may be to develop misgivings about the experimental method yourself. Thus you may end up changing your own attitude as a result of an ingroup member's counterattitudinal behavior.

Three possible mechanisms readily come to mind to explain this reaction to your colleague's action: (a) you may be won over by your colleague's arguments and compelling examples (*persuasion*); (b) you may infer that your colleague must have revised her own attitude, and alter yours to be more in line with hers (*conformity*); or (c) you may spontaneously imagine how uneasy you would feel if in her difficult position, and change your attitude to reduce this discomfort (*vicarious dissonance*). In prior research, we ruled out the persuasion route by showing that attitude change occurs

even when the observer never actually witnesses the actor's persuasive attempts (Norton et al., 2003). The studies presented in this paper test the role of conformity in the attitude change caused by the counterattitudinal behavior of others, in an effort to assert the primacy of the vicarious dissonance model.

Norton et al. (2003) demonstrated the role of vicarious dissonance by using an observer version of the induced compliance paradigm (e.g. Linder, Cooper, & Jones, 1967). In three studies, we found that participants hearing a member of a group with which they strongly identified agree to deliver a counterattitudinal speech changed their own attitude in the direction of the position espoused. We also found that this predicament did not seem to elicit personal discomfort, as personal dissonance does, but instead vicarious discomfort, the discomfort one imagines experiencing in the speaker's place, and that it was this latter form of discomfort that was reduced by attitude change (Elliot & Devine, 1994).

One aspect of vicarious dissonance that makes the phenomenon so intriguing is that it is less directly intuitive than many effects studied in the literature. The reason why individuals change their own attitude in response to an ingroup member's counterattitudinal behavior may not be immediately straightforward. At the individual level, it seems to result from the spontaneous tendency to take the perspective of the member of an important ingroup—participants realize they wouldn't enjoy this situation, and adjust to this by revising their attitude, as they would in a traditional personal dissonance paradigm. At the group level, vicarious dissonance may serve the function of ensuring attitudinal homogeneity (and thus cohesion) of ingroups, even when members have to misrepresent their attitude. Although we have yet to test the situation in which both speaker and observer are naive subjects, the data to date (and the bulk of dissonance research) suggest that vicarious dissonance leads to attitude change in the observer that would mirror that of the actor. The proposition tested in the current paper is that this synchronization is not mediated by a

realization that the actor is changing his or her attitude, but instead by an internalization of the speaker's predicament.

### Goals of the current studies

The primary goal of the current paper is to rule out the alternative conformity interpretation, and to show that attitude change is not the result of perceived attitude change by the speaker and subsequent efforts to conform to that new attitude, but is instead due to the experience of vicarious dissonance. Our previous work (Norton et al., 2003) does not establish whether vicarious dissonance reflects participants' implicit understanding that the speaker experiences cognitive dissonance. Indeed, Bem (1965, 1967) showed that observers of an induced compliance paradigm were able to predict the actor's attitude. The present studies address this issue directly by including a measure of perceived speaker attitude.

To show the limited role of the speaker's perceived attitude, we predict that (a) attitude change should still occur even when perceived speaker attitude is statistically controlled for; (b) making the speaker's antispeech attitude salient should not eliminate attitude change; and (c) standard dissonance manipulations (e.g. aversive consequences) should have a greater impact on participants' attitude than on their perception of the speaker's attitude. Study 1 tests the first and second (as in Norton et al., 2003, Study 2) predictions, while Study 2 tests the first and third.

Finally, the data presented in this paper explore whether our conceptualization of vicarious dissonance as a strict group phenomenon is warranted. The studies conducted thus far have shown that attitude change is greatest when the actor is an ingroup member and when the observer strongly identifies with the group in question. To provide further support for our contention that the phenomenon results primarily from shared group membership, the present studies assess potential individual and interindividual moderators of vicarious dissonance to explore whether they have any impact beyond group affiliation. At

the individual level, we wish to show that unlike group identification, individual differences in empathy (Davis, 1994) do not moderate attitude change. At the interpersonal level, we plan to show that in contrast to group-mediated identification (Hogg, Cooper-Shaw, & Holzworth, 1993), components of interpersonal attraction such as liking and perceived similarity do not moderate attitude change.

### Study 1

In this first study, students who identified strongly with their university heard a fellow student agree to make a speech in favor of giving parents access to students' health records, a measure that was strongly opposed by most students on campus. To test the role of speaker agreement, we explicitly manipulated whether the speaker agreed with the position espoused in the speech, or whether he/she disagreed with it, as in Norton et al.'s (2003) Study 2. Further, to test the importance of personal and interpersonal factors, we measured individual predispositions toward empathy (Davis, 1983) before the experimental session, and liking and perceived similarity at the close of the session. We predicted that attitude change would still occur when the speaker was known to disagree with the speech, that participants would not need to change their perception of the speaker's agreement to change their own attitude, and that attitude change would not be moderated by either a disposition for empathy or personal attraction to the speaker.

### Method

**Pretest** We surveyed our subject pool in a pretest questionnaire about attitudes toward the counterattitudinal issue (giving parents access to students' health records) and identification with the ingroup (Princeton University). The identification scale that we used, adapted from Hogg et al. (1993), consisted of nine items (e.g. *How typical a member of your university do you think you are?*), rated on 7-point scales, which we averaged to create a composite score. We also included a measure of individual differences in

empathy, the Davis Interpersonal Reactivity Index (IRI; Davis, 1983, 1994), which consists of four subscales: *perspective taking*, defined as the tendency to adopt spontaneously the psychological point of view of others in everyday life; *personal distress*, the tendency to experience distress and discomfort in response to extreme distress in others; *empathic concern*, the tendency to experience feelings of sympathy and compassion for unfortunate others; and *fantasy seeking*, the tendency to transpose oneself imaginatively into fictional situations.

**Participants** Fifty-seven Princeton undergraduates took part in Study 1 for payment, and were selected for their opposition to parental access to student health records (a 4 or lower on a 15-point scale), and for high identification with their university (a composite score greater than 4.25, the median of our sample).

**Design** Study 1 used a three-cell design, with a control cell, where the speaker was known to *agree* with the speech that he/she was asked to make, and two experimental cells, where the speaker was known to *disagree*, but with two different orders: in the *disagree/attitude* condition, the participant's attitude was measured before his/her discomfort; in the *disagree/discomfort* condition, it was measured after attitude.

**Procedure** Participants were run in same- and mixed-sex dyads. Two participants were seated in different experimental rooms without ever meeting, under the guise of studying campus 'linguistic subcultures', and were asked to rate the other student's recorded utterances. Participants were always told that the other person, a fellow Princeton student, had been randomly assigned to record a speech, while they would rate that person's speech patterns. The experimenter asked participants to fill out questionnaires for other researchers while he left, ostensibly to record the tape with the other participant, and handed participants the pre-manipulation questionnaire. A few minutes later, he returned with a tape that he played for the participant. Participants listened to a taped

interaction between the speaker and experimenter, and then completed the postmanipulation questionnaire before the actual speech was ever played.

### Materials

**Premanipulation questionnaire** Among unrelated filler surveys, participants in the *agree* condition reported their personal psychological discomfort (items: *uneasy, uncomfortable, bothered*) and positive affect (*happy, good, optimistic*) before hearing the taped interaction. We used 100 mm lines (with endpoints *does not apply at all to applies very much*) and asked participants to mark the point on the line that corresponded to their current affect. This premanipulation affect score gave us a baseline measure with which to compare affect in our experimental cells.

**Tapes** Tapes were prepared with male and female actors, following a scripted interaction reflective of an induced compliance paradigm. On the tape, the experimenter told the speaker that he was combining two projects, the linguistic subcultures project and a project (funded by the Dean's Office) gathering student opinions on a variety of campus topics. He asked the speaker whether he/she would mind making a speech supporting giving parents access to students' health records without prior consent from the students. On the tape, it was made clear that the Dean might use some of the arguments provided by experimental participants to try to implement the unpopular policy. This served to instantiate foreseeable negative consequences (Goethals, Cooper, & Naficy, 1979). Although the speaker was ostensibly given high choice, he/she always agreed to make the speech. The experimenter also asked the speaker whether he/she agreed with giving parental access. In the *agree* condition, the speaker answered 'Actually, I think that's a pretty good idea'; in the *disagree* condition, the speaker answered 'I'd be against that'. This was our manipulation of speaker attitude. The tape was always interrupted before the speech was actually heard, and participants were asked to base their ratings of the speaker's speech patterns on the initial taped interaction.

**Postmanipulation questionnaire** In the *agree* condition, participants reported their attitude toward parental access to health records (on the same 15-point scale used at pretest) immediately following the tape. In the *disagree/attitude* condition, participants heard the tape, then reported their attitude toward the issue, followed by their personal affect. In the *disagree/affect* condition, participants heard the same tape and reported their personal affect, then their attitude. In all conditions, participants then rated the speaker's affect, and their imagined vicarious affect. The affect measures included components of positive affect (items: *happy, good, optimistic*) and psychological discomfort (*uneasy, uncomfortable, bothered*), assessed from three different perspectives: personal affect (. . . *how you feel right now*), the speaker's affect (. . . *how you think the person on the tape was feeling at the end of the taped interaction*), and vicarious affect (. . . *how you think you would have felt at the end of the taped interaction if you had been in the person on the tape's place*). We used 100 mm lines (with endpoints *does not apply at all to applies very much*) and asked participants to mark the point on the line that corresponded to their current affect, the speaker's affect, and their vicarious affect.

Participants then rated their perceptions of the speaker's attitude (on a 15-point scale) and were asked to indicate the speaker's university. They also reported how much they liked the speaker, how similar they thought they were to the speaker, and how typical a member of the university the speaker was, all on 7-point scales (1: *not at all* to 7: *very*).

## Results

**Preliminary analyses** Nine participants were excluded from the analysis because they expressed suspicion, either during the procedure or at debriefing, leaving us with 48 valid participants. Furthermore, analyses of variance (ANOVAs) on the main dependent variables of interest demonstrated that whether the speaker was of the same or of a different gender as the participant did not make a difference. We therefore excluded this factor from the following analyses. We also checked that random

assignment to conditions was successful by verifying that attitudes and identification prior to the manipulation was not significantly different between conditions (both  $F$ s < 1).

**Attitude change** For each participant, we computed an attitude change score by subtracting their attitude score at pretest from their attitude score collected during the experiment. Participants changed their attitude in the two *disagree* conditions, whether the attitude came first ( $M = 0.89$ ,  $SD = 1.6$ ), or second ( $M = 1.31$ ,  $SD = 1.5$ ), whereas there was virtually no change when the speaker agreed with the speech ( $M = 0.15$ ,  $SD = 0.8$ ), replicating the vicarious dissonance effect. The planned contrast comparing the two *disagree* conditions to the *agree* condition was significant ( $t(45) = 2.1$ ,  $p < .05$ ). The average attitude change in the two *disagree* conditions was significantly different from zero ( $t(18) = 2.4$ ,  $p < .03$  and  $t(15) = 3.4$ ,  $p < .005$ , respectively), but not in the *agree* condition ( $t(12) = .7$ , *ns*). Note that, as in previous uses of this technique (Elliot & Devine, 1994), counterbalancing the order of questions in the two *disagree* conditions did not impact attitude change ( $t(33) = -.78$ , *ns*).

**Speaker agreement** Participants in the *agree* condition rated the speaker as being more in favor of parental access to health records ( $M = 13.15$ ) than those in the *disagree/attitude* ( $M = 1.84$ ) and *disagree/affect* ( $M = 2.44$ ) conditions ( $F(2, 45) = 481.24$ ,  $p < .001$ ), reflecting the success of our manipulations. More importantly, these means show that participants who witness speakers disagree with the issue and then agree to make a speech in favor do not infer that speakers have switched to agreeing with the issue. Furthermore, above we observed most attitude change in the *disagree* conditions, whereas the speaker was perceived as most supportive in the *agree* condition.

To test our main prediction that perceived speaker attitude cannot account for the observed attitude change, we conducted an analysis of covariance (ANCOVA) on attitude change with agreement as a factor and speaker's attitude as a covariate. Although

speaker's attitude was a marginally significant covariate ( $F(1,45) = 3.1, p < .09$ ), it failed to account for the effect of the agreement manipulation, which was still significant ( $F(1, 45) = 4.7, p < .04$ ). Thus attitude change does not seem to be the result of changing one's perception of the speaker's attitude. Between conditions, perceived speaker attitude seemed indeed out of step with attitude change, as reflected by an overall correlation coefficient of  $-.23, ns$ . Within conditions, however, it appears that the correlation is nearly significant and positive in the disagree condition ( $r = .33, p < .06$ ), whereas it is smaller and negative in the agree condition ( $r = -.27, ns$ ).

**Individual differences in empathy** We computed Davis' IRI subscales for each participant from the pretest measures, yielding four scores: *fantasy seeking* (FS), *personal distress* (PD), *perspective taking* (PT), and *empathic concern* (EC). Only FS and PD correlated significantly ( $r = .30, p < .05$ ). First, none of these subscales correlated with attitude change. Second, using median splits on each of the four dimensions, we ran four new ANOVAs on the attitude change scores, using manipulated speaker agreement and level of empathy on each of the four dimensions. In none of the analyses did either the main effect for empathy subscale or the interaction term come close to significance (all  $F_s < 1$ ), whereas speaker agreement remained significant.

**Similarity, liking, and typicality** Speakers who disagreed were perceived as more similar ( $M = 3.9, SD = 1.4$ ) than those who agreed ( $M = 2.9, SD = 1.4$ ) ( $t(46) = -2.24, p < .04$ ). However, they were liked as much ( $M = 4.1, SD = 1.3$  vs.  $M = 4.0, SD = 1.3$ ) ( $t(46) = .25, ns$ ), and seen as similarly typical ( $M = 5.2, SD = 0.9$  vs.  $M = 4.9, SD = 1.0$ ) ( $t(46) = .34, ns$ ). None of these measures was correlated with attitude change (similarity:  $r = -.04$ ; liking:  $r = -.10$ ; typicality  $r = .03$ ; all  $ns$ ). Furthermore, after using median splits, we conducted three ANOVAs on the attitude change scores, using manipulated speaker agreement and level of liking, similarity, or typicality as factors. In all three cases,

neither the main effect for the additional variable nor the interaction approached significance ( $F_s < 1$ ).

**Affect measures** We created six affect indices by aggregating ratings of the different trait states by perspective and valence: personal positive affect (P+), other's positive affect (O+), vicarious positive affect (V+), personal psychological discomfort (P-), other's psychological discomfort (O-), and vicarious psychological discomfort (V-). As with attitude change, none of the affect measures differed significantly due to order in the disagree conditions. Given this lack of order effect, all the analyses below are collapsed, comparing the agree condition to the two disagree conditions. The speaker agreement manipulation had an impact on all of the affect variables. Overall, positive affect was higher in the agree conditions ( $M_s = 68, 46, \text{ and } 42$  for P+, V+, and O+) than in the disagree conditions ( $M_s = 50, 33, \text{ and } 31; p < .05$  for P+,  $p < .10$  for the other two). Psychological discomfort, on the other hand, was lower in the agree condition ( $M_s = 12, 39, \text{ and } 54$  for P-, V-, and O-) than in the disagree condition ( $M_s = 26, 65, \text{ and } 69; p < .005$  for V-,  $p < .05$  for the other two). Of these six measures, only V- was significantly correlated with attitude change ( $r = .31, p < .05$ ). The scale that otherwise correlates most is P- ( $r = .26, p < .09$ ), while the other correlations do not approach significance. We also correlated the six affect scores with the four components of the Davis IRI. The only significant correlation was between V- and empathic concern,  $r = .29, p < .05$ . In other words, the degree to which our participants were dispositionally inclined to empathize with others was related to the amount of vicarious negative affect they reported.

### Discussion

In this first study, participants expressed more support for an unpopular policy after they heard a fellow student agree to make a speech in favor of it, but only if they knew that the speaker had the same misgivings about the policy as they did. Not only did they not show less attitude change when the speaker

expressed his or her disagreement, ruling out the conformity interpretation, but in fact they showed *more* attitude change when the speaker disagreed than when she agreed, replicating the pattern observed in Norton et al.'s Study 2 (2003). Defendants of the conformity interpretation could argue that participants disregarded this initial statement and assumed that the speaker's attitude was in line with his or her speech. To counter this possibility, we also collected perceived speaker attitude toward the end of the session, and found that in the disagree conditions, speakers were still perceived as clearly disagreeing with the position espoused, and yet these are the conditions that yielded most attitude change. Most importantly, the difference in means remained significant once perceived speaker attitude was controlled for. This pattern of results provides strong support for the vicarious dissonance position against a conformity interpretation of our prior results. It is when the speaker was known to disagree with the position he or she was willing to take that participants changed their attitudes most. However, this attitude change does not seem to result from a perception that the speaker changed his or her own attitude. An examination of the affect measures suggests that participants change their attitudes to reduce their vicarious discomfort, the unease they spontaneously imagine experiencing had they been in the speaker's shoes.

Study 1 also demonstrates that individual and interindividual factors have little impact beyond affiliation on the attitude change observed, supporting the view that this phenomenon is primarily a product of individuals' group memberships. First, among our high affiliators, individual differences in empathy did not correlate with attitude change, nor did they moderate the effect of speaker agreement. Second, interindividual factors (liking for the speaker and perceived similarity) did not correlate with attitude change or moderate the impact of speaker agreement either. One might wonder whether the lack of predictive value of individual empathy could be attributed to the quality of the measure or the delay between the measure and the experiment.

Although this is a valid concern, the predictive value of the scale is supported by the fact that empathic concern was indeed correlated significantly with vicarious discomfort ( $r = .29$ ,  $p < .05$ ). Individuals who were predisposed to empathically connect with others in distress reported higher levels of negative vicarious affect.<sup>1</sup> Further, by suggesting what type of person might be more likely to experience vicarious discomfort, it provides us some insight about the nature of vicarious discomfort.

It should be noted that by selecting high affiliators, we are working with a truncated range of participants. It is quite possible that, had we included low affiliators in our study, the factors listed above would have moderated the effect. What the current results suggest is that these factors do not play a role within our high affiliators, suggesting that they would not play a role once affiliation is controlled for in a fuller range, and thus that they do not play a role above and beyond that of affiliation. It is possible, however, that among low affiliators these personality variables play more of a role in moderating vicarious processes; future research should explore this possibility.

The absence of an order effect on the affect measure, which has in the past been taken as evidence that attitude change reduces dissonance (see Elliot & Devine, 1994, and Norton et al., 2003, Study 3), can be attributed straightforwardly to a change in the procedure. Whereas Norton et al. (2003) counterbalanced the measure of attitude and measures of affect from all three perspectives (self, other, vicarious), in the current study only personal affect was counterbalanced with attitude, and vicarious attitude, which was shown to be reduced in Norton et al.'s Study 3 (2003), was always measured last. Therefore we wouldn't have expected it to be influenced by order.

Note that, as previous dissonance investigators, we did not predict a mediating role for discomfort. The link between attitude change and negative arousal or psychological discomfort in the dissonance literature is complex and cannot unfortunately be reduced to a simple mediation analysis. A quick review of the recent dissonance literature illustrates the thorniness

of this issue: Elliot & Devine (1994) predicted and found a negative correlation between discomfort and attitude change in their key dissonance cells when discomfort is measured after attitude (e.g.  $r = -.22$ ,  $p = .13$ , one-tailed in their Study 2), as did Losch and Cacioppo (1990) when they measured skin conductance before attitude ( $r = -.65$ ,  $p < .01$ ). On the other hand, Harmon-Jones, Brehm, Greenberg, Simon, and Nelson (1996) report a positive correlation ( $r = .42$ ,  $p < .05$ ) between skin conductance and attitude change across conditions when conductance is measured before attitude. The source of the complexity is that the model is 'hydraulic' or homeostatic, so the prediction is that the very fact of changing one's attitude could reduce some discomfort (personal or vicarious). Whereas greater discomfort should lead to more attitude change (the traditional second branch of the indirect path in a mediation analysis), more attitude change should in turn reduce discomfort. In other words the model predicts a positive influence in one direction and a negative one in the other; the two influences could thus cancel each other out, or the correlation could go either way. Indeed, although we do not find a mean difference on V- depending on when it is assessed, we do find a marginal positive correlation in the cell where affect was measured first ( $r = .43$ ,  $p < .10$ , two-tailed), but none in the other two cells ( $p > .45$  in both), precluding a mediation analysis.

Study 1 was able to show that attitude change occurs even when it is clear that the speaker disagrees with the speech, and yet, like Norton et al.'s Study 2 (2003), these results may not be enough to rule out the conformity interpretation altogether. To a large extent, showing that participants perceived more speaker disagreement in the speaker *disagree* condition boils down to a check of the speaker agreement manipulation. We were heartened to find out that participants in the *disagree* condition thought the speaker disagreed with the issue even after he or she agreed to make the speech. However, it is actually still possible that participants may have perceived attitude change in the speaker as a result of agreeing to make the speech, as in Bem's (1965, 1967) interpersonal replications of

dissonance paradigms: although the average perceived speaker attitude in the *disagree* conditions was 2.1 on a 15-point scale after they heard the tape, it is possible that it would have been significantly lower had it been measured before the tape, and that participants were indeed inferring attitude change in the speaker, as predicted by the conformity model. Note that the participants' own attitude in the *disagree* cells went from 1.5 at pretest to only 2.6 after the manipulation. Thus it seems necessary to add a relevant comparison condition, in which the speaker attitude would not be expected to change, to compare it with perceived speaker attitude in the vicarious dissonance cells. Study 2 accomplishes this by manipulating the presence of foreseeable negative consequences.

## Study 2

Study 1 replicates the vicarious dissonance effect and casts doubt on a conformity interpretation of the phenomenon. However, in Study 1 because the vicarious dissonance manipulation *centered on* speaker agreement, it precludes testing whether the conditions of vicarious dissonance lead to a change in perceived speaker attitude. Study 2 enables us to do this by having participants listen to a speaker who always explicitly disagrees with the position espoused in the speech, while manipulating in a simple two-group design the presence or absence of foreseeable aversive consequences (Cooper & Fazio, 1984). We predict that the presence of aversive consequences will lead to attitude change in our participants, but that such consequences will not lead to greater perceived agreement by the speakers. Further, we again predict no correlation between our participants' own attitudes and the attitude they attribute to the speaker. Study 2 again includes measures of individual differences in empathy as well as liking and similarity, to provide additional tests of ancillary predictions.

## Method

**Pretest and participants** We used the same pretest survey as in Study 1, measuring group identification and campus attitudes. Twenty-five

Princeton undergraduates took part in Study 2 for experimental credit or payment, and were selected on the same criteria as Study 1.

**Procedure** The procedure was similar to that of Study 1, but we changed the taped interaction in two significant ways. First, all participants were led to believe that the speaker disagreed with the speech. Second, the presence of aversive consequences was manipulated in the interaction between the speaker and experimenter. In the *aversive consequences* condition, the experimenter explained that the Dean needed students to make speeches in favor of letting the health center call parents about students' medical problems so that the Dean could use the arguments generated to justify such a policy change, which served to instantiate aversive consequences. In the *no consequences* condition, following the procedure used by Goethals et al. (1979), the dialogue was the same, except that the speaker was informed that the tape would be reviewed only by the experimenter, would be completely confidential, and would be erased after the session. All speakers then agreed to make a speech in favor (though as in Study 1, the speech was not actually played). Participants were run in same- and mixed-sex dyads.

### Results

**Preliminary analyses** No participant reported suspicion in Study 2.<sup>2</sup> As before, preliminary analyses revealed no main effects or significant interactions for gender of partner, so we do not report further analyses for this factor. We found no significant difference in premanipulation measures of attitude and identification (both  $t(23) < 1.2$ , *ns*), suggesting that random assignment to conditions was successful on those dimensions.

**Manipulation checks** As expected, participants in the aversive consequences condition said they would have been less willing to make the speech themselves ( $M = 3.54$ ) than those in the no consequences condition ( $M = 4.75$ ) ( $t(23) = 1.7$ ,  $p < .05$ , one-tailed). All participants correctly identified the speaker's university.

**Attitude change** Attitude change scores were computed as before. As predicted, participants in the *no consequences* condition did not change their attitudes ( $M = 0.33$ ,  $SD = 0.65$ , which is not different from zero,  $t(11) = 1.8$ , *ns*), whereas those in the *consequences* condition did ( $M = 1.5$ ,  $SD = 1.7$ ,  $t(12) = 3.2$ ,  $p < .01$ ). The difference between the two conditions was also significant ( $t(23) = -2.3$ ,  $p < .05$ ).

**Attitude of speaker** We did not find differences in the attitude ascribed to the speaker depending on the presence of negative consequences: the speaker was seen to disagree strongly both in the *consequences* condition ( $M = 1.9$ ,  $SD = 1.1$ ) and in the *no consequences* condition ( $M = 2.4$ ,  $SD = 1.0$ ), and this difference was not significant ( $t(23) = 1.4$ , *ns*). It should be noted that to test the difference in attitude ascribed to the other, we did not benefit from the added power afforded by the pretest attitude measure that we used to compute an attitude change score in the case of personal attitude above. However, the results do not suggest that more power would show a similar pattern in perceptions of the other as in the self: closer inspections of the means above reveal that whereas personal attitude is higher in the consequences condition as expected, attitude ascribed to the other is actually *lower* (though, again, not significantly so) with consequences than without. When we conducted a  $2 \times 2$  mixed model ANCOVA with condition as a between-participants factor, other vs. own attitude as a within-participants factor, and pretest attitude as a covariate, we found a highly significant interaction between condition and target ( $F(1,22) = 11.1$ ,  $p < .005$ ), suggesting that the impact of the manipulation on one's own attitude was far greater than it was on the perception of the speaker's attitude. The consequences manipulation affected the attitude attributed to the other in the opposite way that it was influencing own attitude. Thus it is difficult to interpret the vicarious dissonance effect by assuming that it results from our participants perceiving attitude change in the other.

Furthermore, to test as in Study 1 whether perceived speaker attitude accounts for attitude

change, we conducted an ANCOVA with speaker agreement as a covariate, and condition otherwise predicting attitude change. We found that speaker agreement was indeed a significant covariate ( $F(1,22) = 7.0, p < .02$ ), but that it didn't account for the impact of condition on attitude change, which remained significant ( $F(1,22) = 10.2, p < .005$ ). This suggests that the attitude change effect is not a result of a change in the perceived speaker attitude. Speaker attitude was correlated with attitude change ( $r = .31, p = .13$ ), especially within conditions ( $r = .50, p < .08$  with consequences,  $r = .61, p < .04$  without). These correlations may be traced to a link between own and perceived speaker attitude (zero-order  $r = .37, p = .07$ ;  $r = .43, p = .03$  when condition is partialled out). In retrospect, this positive correlation should have been predicted given the vast literature on social projection (e.g. Allport, 1924; Monin & Norton, 2003). Indeed, a re-examination of Study 1 shows that although the overall correlation between own and perceived speaker attitudes was not significant ( $r = -.18$ ), in the *disagree* conditions (where participants felt similar to the speaker) the correlation was positive and significant ( $r = .34, p < .05$ ) whereas it was  $-.16, ns$ , in the *agree* condition. Again though, the important finding for our purpose is that controlling for perceived speaker attitude does not eradicate the impact of the manipulation on attitude change.

**Individual differences in empathy** When we looked at the IRI subscales, FS was correlated with PT ( $r = .63, p < .01$ ) and EC ( $r = .47, p < .05$ ), and PD was correlated with EC ( $r = .40, p < .05$ ). First, none of these individual difference measures correlated significantly with attitude change. Second, when we ran four separate  $2 \times 2$  ANOVAs with consequences and level of empathy subscale as factors, we found as before that none of the subscales came close to interacting with the manipulation, although this time there was a marginal main effect for personal distress ( $F(1, 21) = 4.0, p < .06$ ), such that participants more prone to personal distress showed more attitude change ( $M = 1.4$  vs.  $0.4$ ). This echoes the marginal positive

correlation ( $r = .34, p < .10$ ) found between disposition for personal distress and attitude change across conditions. The important point for our purpose is the lack of interaction between any of these subscales of empathy and the consequences manipulation.

**Similarity, liking, and typicality** Participants in the aversive consequences condition did not see speakers as less similar to themselves than participants in the no consequences condition ( $M = 3.9, SD = 1.1$  vs.  $M = 3.8, SD = 1.7$ , respectively;  $t(23) = -.02, ns$ ) nor as less likable ( $M = 4.4, SD = 0.8$  vs.  $M = 4.1, SD = 1.3$ ;  $t(23) = -.71, ns$ ) nor as less typical ( $M = 5.5, SD = 1.1$  vs.  $M = 5.0, SD = 1.3$ ;  $t(23) = -1.12, ns$ ). Similarity and liking did not correlate with attitude change ( $r = .02$  and  $.12$ , respectively, both *ns*), but typicality showed a marginal positive trend ( $r = .37, p = .07$ ), such that speakers seen as more typical led to more attitude change. Using median splits, we ran three ANOVAs predicting attitude change with consequences and level of liking, similarity or typicality as factors. Liking, similarity, and typicality did not impact attitude change significantly, neither as a main effect nor in interaction with aversive consequences (all *F*s *ns*).

**Affect measures** In contrast with Study 1, P+ and P- did not differ by condition, nor did O+ and O-, with only O+ approaching significance ( $t(23) = 1.8, p < .08$ ). However, as predicted, V- and V+ were markedly different in the consequences condition ( $M$ s = 65 and 34, respectively) and the no consequences condition ( $M$ s = 41 and 51) ( $t(23) = -3.7$  and  $2.5$ , respectively, both  $p$ s  $< .05$ ). In sum, our participants experienced more vicarious discomfort when the speech had aversive consequences than when it did not, paralleling our results for attitude change. As before, attitude change correlated the most with V- ( $r = .24$ ) and V+ ( $r = -.21$ ), but this time neither reached conventional levels of significance (both  $p$ s  $> .25$ ), maybe because of the low power afforded by the small number of subjects. Discomfort measures were also correlated with the individual difference empathy subscales. P+ correlated with PT ( $r =$

.46,  $p < .05$ ), and as in Study 1,  $V-$  correlated with EC ( $r = .47$ ,  $p < .05$ ).

### Discussion

When high identifiers heard a fellow student agree to make a speech in favor of a position with which both of them disagreed, one that had concrete negative consequences, they changed their attitude toward the position the speaker agreed to take in a speech, even when participants did not hear the speech itself. Those participants who believed that the speech was going to be used against students by a university dean expressed the highest levels of vicarious discomfort and changed their attitude most, mirroring the moderating effect of aversive consequences observed in personal dissonance (Cooper & Fazio, 1984).

More important, this study goes beyond the manipulation of speaker agreement in Study 1 to show that when vicarious dissonance is manipulated through aversive consequences, observers' attitude change occurs even in the absence of any corresponding change in the perceived attitude of the speaker. Along with the fact that vicarious discomfort was significantly higher in the aversive consequences conditions whereas perceived speaker discomfort was not affected significantly, this finding further supports the interpretation that vicarious dissonance occurs when participants spontaneously take the perspective of the member of an important group and imagine how they would feel in their place, rather than as a result of trying to infer the group member's attitude about the topic in question and changing their own attitudes to correspond. In this study, vicarious discomfort was always measured after attitude, precluding a standard mediation analysis given the issues raised in the discussion to Study 1; indeed, the within-cell correlations between attitude change and vicarious discomfort were not significant (both  $ps > .40$ ).

### General discussion

When we witness a member of an important ingroup engage in counterattitudinal behavior, we change our attitude in the direction of her

behavior. Two studies demonstrate this phenomenon, and show that it is unlikely to be the result of inferring that the speaker herself changed her attitude to bring it more in line with her behavior. In Study 1, participants showed more attitude change when they knew speakers disagreed with the speech they were required to make. In Study 2, aversive consequences brought about attitude change in the observer but did not impact perceptions of the speaker's attitude. In both studies, analyses of covariance controlling for perceived speaker attitude did not eliminate the effects observed. Taken together, these studies suggest that conformity to the speaker's attitude is unlikely to play a causal role in the attitude change observed in the vicarious dissonance paradigm. These studies also further rule out a role for persuasion in accounting for attitude change, because the effect occurs despite the fact that the actor's actual persuasive arguments were never heard. It seems to be enough to hear the ingroup member agree to engage in the counterattitudinal behavior for our participants to start imagining how uneasy they would be in that same situation and change their own attitudes as a result.

It is possible that the lack of impact of our vicarious dissonance manipulations on perceived speaker attitude (despite a clear predicted effect on own attitude) might have to do with the placement of this variable, which was always measured after participants reported their own attitudes. We see this possibility as somewhat unlikely. These types of order effects, reflective of a hydraulic model, are most likely to be found in dissonance paradigms when measures of discomfort are given after an attitude measure (e.g. Elliot & Devine, 1994; Norton et al., 2003), but measures of constructs that are less 'on-line' than affect or discomfort are probably less likely to be influenced. Attitude measures, for example, are typically not influenced by order in these paradigms, and if vicarious dissonance had led to perceived attitude change in the speaker or distancing from the speaker, it seems it should have resulted in a general different impression of the speaker that should have subsisted even if the relevant

question is asked sixth or seventh on the same page. However, the possibility remains, and in future studies it would be useful to measure perceived speaker attitude just after the end of the taped interaction.

***The lack of impact of individual and interindividual factors***

Furthermore, these studies showed that individual predispositions toward empathy, as measured by Davis' IRI scale, as well as interindividual idiosyncrasies, such as liking for the speaker and perceived similarity with the speaker, did not impact attitude change, nor did they interact with the active manipulations. Previous studies have shown instead that group identification was a strong moderator of attitude change in this paradigm. This suggests that vicarious dissonance is not attributable to a subset of particularly empathetic people, and has more to do with affiliation with a common group identity than affiliation with any one particular group member, a distinction highlighted in recent years by Hogg and his colleagues (e.g. Hogg et al., 1993). The possibility remains, however, that vicarious dissonance could be experienced in particularly close dyads that have become part of one's own identity, as when parents observe children (Barquissau, Schmeder, & Lickel, 2004) or spouses observe each other acting counter-attitudinally.

Two possible concerns should be addressed here. First, whereas group identification and individual empathy were always measured in a separate prior session, the measures of liking and perceived similarity were collected at the end of the experimental session, raising the possibility that the consequences manipulation may have influenced them. This was necessary since it was impossible to have participants rate their liking for the target before witnessing the target during the experiment. However, this concern is reduced by the lack of mean differences on these measures, as tested prior to conducting the internal analyses using them as factors. Second, because the current studies were conducted only with high identifiers, one may wonder whether a restriction of range on

the empathy measures may have limited the capacity to test their impact. This assumes that there might be a relationship between empathic disposition and ingroup identification. Future studies should include high and low identifiers and include measures of empathy to ascertain that ingroup identification is indeed the only individual difference that impacts attitude change. The current studies thus provide a weak test of the zero-order impact of these variables on attitude change, but they do provide a good test of their impact above and beyond the role of identification; in that respect the results suggest that their variables provide little additional explanatory power. One possibility that would make this interpretation problematic is if these variables had a larger impact among low identifiers; future research should explore this possibility.

In conclusion, the power of our group identifications extends where we least expect it. Identifying strongly with a group not only means that we will invest time and energy in that group, seek out group activities, or derive much of our daily satisfaction from that group, but also means that we are implicated in the behavior of the members of that group. Thus, paradoxically, upon merely witnessing a fellow group member doing something with which we do not quite agree, we will change our own attitude to be more in line with their behavior. Of course, changing our own attitudes also serves to make the group member's initially surprising behavior somewhat more acceptable, and thus vicarious dissonance may in part serve an important function in preserving group cohesiveness in the context of discrepant behavior: not only is the phenomenon the result of ingroup identification, but it can, in the long run, contribute to the well-being of that group. This research also sheds insight into the nature of ingroup identification itself. These results, coupled with the results in Norton et al. (2003), suggest that the extension of the self provided by group membership may lead to spontaneous identification with ingroup members, particularly in situations that may be stressful for them. A fascinating avenue for future research would

be to test whether other classic individual social psychological phenomena occur vicariously for high identifiers in the presence of an ingroup member. It may turn out that we live much more vicariously than we ever imagined.

## Notes

1. Although the lack of correlation between vicarious discomfort and the perspective taking (PT) dimension of the IRI scale may seem at first glance surprising, it is important to keep in mind that the items comprising the PT subscale have more to do with taking the point of view of the other side in disagreements (e.g. 'I believe there are two sides to every question and try to look at them both') than with the capacity to imagine how it would *feel* to be in the other's shoes, which is better captured by the items of the empathic concern (EC) subscale (e.g. 'I often have tender, concerned feelings for people less fortunate than me', 'Other people's misfortunes do not disturb me a great deal' (reversed), or 'When I see someone being taken advantage of, I feel kind of protective toward them'). Similarly, EC describes interpersonal emotions that seem to capture better the experience of V- than the scale of personal distress (PD) which is more self-centered (e.g. 'Being in a tense emotional situation scares me' or 'I tend to lose control during emergencies'). It therefore seems quite appropriate that EC would be the scale that most correlates with V-.
2. A majority of suspicious participants in Study 1 were in the *agreement* condition, which may have sounded odd because the speaker espoused a position that was unexpected for a student (basically agreeing to a violation of his/her privacy)—and this condition was absent from Study 2.

## References

- Allport, F. H. (1924). *Social psychology*. New York: Houghton Mifflin.
- Barquissau, M., Schmeder, T., & Lickel, B. (2004, February). *Growing pains: Parents' experience of vicarious shame and guilt*. Poster session presented at the Annual Meeting of the Society of Personality and Social Psychology, Austin, TX.
- Bem, D. J. (1965). An experimental analysis of self-persuasion. *Journal of Experimental Social Psychology*, *1*, 199–218.
- Bem, D. J. (1967). Self-perception: An alternative explanation of cognitive dissonance phenomena. *Psychological Review*, *74*, 183–200.
- Cooper, J., & Fazio, R. H. (1984). A new look at dissonance theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 17). New York: Academic Press.
- Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, *44*, 113–126.
- Davis, M. H. (1994). *Empathy: A social psychological approach*. Boulder, CO: Westview Press.
- Elliot, A. J., & Devine, P. G. (1994). On the motivational nature of cognitive dissonance: Dissonance as psychological discomfort. *Journal of Personality and Social Psychology*, *67*, 382–394.
- Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.
- Gilbert, D. T., & Malone, P. S. (1995). The correspondence bias. *Psychological Bulletin*, *117*, 21–38.
- Goethals, G. R., Cooper, J., & Naficy, A. (1979). Role of foreseen, foreseeable, and unforeseen behavioral consequences in the arousal of cognitive dissonance. *Journal of Personality and Social Psychology*, *37*, 1179–1185.
- Harmon-Jones, E., Brehm, J. W., Greenberg, J., Simon, L., & Nelson, D. E. (1996). Evidence that the production of aversive consequences is not necessary to create cognitive dissonance. *Journal of Personality and Social Psychology*, *70*, 5–16.
- Harmon-Jones, E., & Mills, J. (Eds.) (1999). *Cognitive dissonance: Progress on a pivotal theory in social psychology*. Washington, DC: American Psychological Association.
- Hogg, M. A., Cooper-Shaw, L., & Holzworth, D. W. (1993). Group prototypicality and depersonalized attraction in small interactive groups. *Personality and Social Psychology Bulletin*, *19*, 452–465.
- Kelley, H. H. (1967). Attribution theory in social psychology. *Nebraska symposium on motivation*, *14*, 192–241.
- Linder, D. E., Cooper, J., & Jones, E. E. (1967). Decision freedom as a determinant of the role of incentive magnitude in attitude change. *Journal of Personality and Social Psychology*, *6*, 245–254.
- Losch, M. E. & Cacioppo, J. T. (1990). Cognitive dissonance may enhance sympathetic tonus, but attitudes are changed to reduce negative affect rather than arousal. *Journal of Experimental Social Psychology*, *26*, 289–304.
- Monin, B., & Norton, M. I. (2003). Perceptions of a fluid consensus: Uniqueness bias, false consensus.

false polarization and pluralistic ignorance in a water conservation crisis. *Personality and Social Psychology Bulletin*, 29, 559–567.

Nisbett, R. E., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice Hall.

Norton, M. I., Monin, B., Cooper, J., & Hogg, M. A. (2003). Vicarious dissonance: Attitude change from the inconsistency of others. *Journal of Personality and Social Psychology*, 85, 47–62.

Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 174–221). New York: Academic Press.

*Paper received 21 July 2003; revised version accepted 1 April 2003.*

### **Biographical notes**

BENOÎT MONIN (PhD) is an assistant professor of psychology at Stanford University. His research focuses on social norms, how we perceive them, what we do to avoid them, and how we react to those who break them.

MICHAEL I. NORTON (PhD) is a postdoctoral fellow at the Sloan School of Management at the Massachusetts Institute of Technology. His current interests include self-deception, vicarious processes, and political correctness.

JOEL COOPER (PhD) is a professor of psychology at Princeton University. His research interests include attitudes and attitude change (particularly as they relate to the process of cognitive dissonance), the effect of expert testimony in courts of law, and gender differences in the effectiveness of information technology, particularly among schoolchildren.

MICHAEL HOGG (PhD) is an Australian professorial fellow at the University of Queensland and a visiting professor of psychology at the University of California, Santa Barbara. He is a Fellow of the Academy of the Social Sciences in Australia and a Fellow of the Society for the Psychological Study of Social Issues. His research focuses on group processes, intergroup relations, and self-conception, and has been closely associated with the social identity perspective.