Keeping it in-house: How audience affects responses to group criticism

MATTHEW J. HORNSEY\textsuperscript{1*}, PETRA DE BRUIJN\textsuperscript{2}, JENNIFER CREED\textsuperscript{1}, JAMES ALLEN\textsuperscript{1}, AMARINA ARIYANTO\textsuperscript{3} AND ALICIA SVENSSON\textsuperscript{1}

\textsuperscript{1}University of Queensland, Australia
\textsuperscript{2}University of Amsterdam, The Netherlands
\textsuperscript{3}University of Indonesia, Indonesia

Abstract

It is often assumed that group-directed criticism is best kept ‘in-house’, but the effects of audience on responses to criticism have not been directly examined. Consistent with predictions, ingroup members who criticized the group to an outgroup audience were seen to be making a less appropriate choice of audience (Experiments 2 and 3), aroused more negative feelings (Experiment 1), were downgraded more strongly on personality traits (Experiment 2), and were seen to be doing more damage to the group (Experiment 2) than were ingroup members who kept their criticisms in-house. Experiment 3 showed that, whereas moderate identifiers agreed with the comments less and showed weaker friendly intentions toward the critic when an outgroup audience as compared to an ingroup audience was chosen, high identifiers agreed with the criticisms just as strongly—and showed more friendly intentions toward the critic—when they were made to an outgroup as compared to an ingroup audience. Results are discussed in light of the broader literature on identity threat. Copyright © 2005 John Wiley & Sons, Ltd.

There are times when people feel compelled to stand up and articulate their group’s shortcomings, an act that carries with it enormous social risks. Research on dissent (e.g. Festinger, 1950, 1954) shows that people who do not conform to the ‘party line’ face social censure, and the black sheep literature shows how quick groups are to distance themselves from members who reflect poorly on the group (e.g. Marques & Paez, 1994). Indeed, a mechanistic reading of social identity theory (Tajfel & Turner, 1979; Turner, 1999) might lead one to believe that ingroup critics are doomed to face defensiveness because they are attacking a part of people’s self-concept.

In many cases, however, an ingroup critic is doing more than attacking the group—they are trying to incite positive change. Criticism might make group members feel uncomfortable, but it can be a crucial

\*Correspondence to: Dr M. J. Hornsey, School of Psychology, University of Queensland, St Lucia 4072, Queensland, Australia.
E-mail: m.hornsey@psy.uq.edu.au

Contract/grant sponsor: Australian Research Council-Discovery.

Copyright © 2005 John Wiley & Sons, Ltd.

Received 13 August 2004
Accepted 12 September 2004
pre-requisite for the re-invigoration, re-invention, and reform of a group culture. Furthermore, legitimate criticism of a group’s direction can be stimulating in terms of promoting innovation, creativity, and flexibility in decision making (Janis, 1982; Nemeth & Owens, 1996). Indeed, defensiveness in the face of legitimate criticism can have disastrous consequences for a group in the sense that it can allow maladaptive or even immoral elements of a group’s culture to survive. So the ingroup critic is likely to arouse a complex set of feelings. On the one hand they are potentially damaging the integrity of an identity that is important to members’ self-definition. On the other hand they might be seen as doing long-term good to the group; as unpleasant but necessary, much like medicine.

Indeed, previous research on group criticism demonstrates surprisingly high levels of generosity toward ingroup critics. Hornsey, Oppes, and Svensson (2002, Study 1a) showed Australians scripts in which targets criticized Australia. These comments were then attributed to either an Australian (an ingroup critic) or a non-Australian (an outgroup critic). They found that criticisms aroused less defensiveness when they were made by an ingroup member than when the very same comments were made by an outgroup member (‘it’s OK if we say it, but you can’t’). This effect was later replicated using other identities such as ‘university student’ and ‘maths-science student’: ingroup critics were seen to be more likable, aroused less sensitivity, and were agreed with more strongly than were outsiders who made the same comments (Hornsey et al., 2002, Studies 1b and 2). This tendency—labelled the intergroup sensitivity effect—survived even when the outsider had as much experience with the group as ingroup members (Hornsey & Imani, 2004).

Drawing on research on persuasion (e.g. Eagly, Wood, & Chaiken, 1978; Wood & Eagly, 1981), Hornsey and colleagues (Hornsey & Imani, 2004; Hornsey, Trembath, & Gunthorpe, 2004) argued that, when deciding how to respond to group-directed criticism, group members engage in a process of hypothesis testing about motive (‘Why are they saying this?’). Consistent with social identity research on trust (e.g. Brewer, 1981; Vivian & Berkowitz, 1992, 1993; Worchel, 1979) it was argued that group membership operates as a heuristic that helps gauge whether a critic’s motives are honourable or not. When criticisms stem from an insider, people are more likely to make the attribution that the criticisms are made with the best interests of the group at heart than when the same comments are made by an outsider. Across three studies, Hornsey and Imani (2004) found some support for this identity-based explanation of the intergroup sensitivity effect. Ingroup critics were attributed more constructive motives for their comments than were outgroup critics, and this attributional bias mediated the effects of group membership on ratings of personality evaluations, negativity toward the criticism, and agreement with the comments.

In summary, Hornsey and colleagues (Hornsey et al., 2002; Hornsey & Imani, 2004) showed that the motives of ingroup critics are generally trusted and that this helps ease resistance to criticism. But what if group members have cause to doubt the commitment of the critic to the group? In these circumstances, one might expect that the generous attributions typically made of the ingroup critic might be withdrawn and the ingroup critic might face heightened levels of resistance. An experiment by Hornsey et al. (2004) provides some evidence for this. Insiders who criticized their group aroused less defensiveness than did outsiders, but only when the insider was described as having an ongoing commitment to the group. When the insider was described as a low identifier they aroused as much defensiveness as did the outsider.

AUDIENCE AND CRITICISM

In the previous section we have seen how cues about whether an ingroup member cares about the group can be crucial in determining whether or not people are prepared to accept the criticisms they make. What is missing from this research, though, is an appreciation of audience and how the group
membership of the audience can influence how criticisms are received. In the traditional paradigm, the
interviewees are ostensibly making their comments to an experimenter, presumably because they were
requested to do so. In real life, though, criticisms are directed to an audience, and the choice of
audience might play a large role in influencing how people respond. For example, it seems that
Australians are relatively tolerant of another Australian who criticizes the group in an interview with
an experimenter (Hornsey & Imani, 2004). But what if the person was to make these comments to a
group of Americans, or Brits? If an ingroup member criticizes their group to an outgroup audience,
does this increase defensiveness?

There are at least two reasons to believe that it would. First, an ingroup member who makes their
comments to an outgroup audience might invite suspicion with regard to their motives. For example,
the choice of an outgroup audience might raise suspicion that the ingroup member is planning to shift
allegiances and that their act of criticism is an attempt to distance themselves from their old group and
to ingratiate themselves with the group to which they intend to migrate (see Eagly et al., 1978, for a
related argument in the domain of persuasion). Alternatively, it could be that the ingroup critic has
psychologically withdrawn from the ingroup and seeks the approval of an outgroup to which they feel
more psychologically aligned. Either way, the choice of an outgroup audience might arouse suspicion
about commitment and motive, which we have already seen can have implications for whether the
critic faces defensiveness.

A second reason why ingroup critics might be downgraded for taking their comments to an
outgroup audience is that the critic might be seen to be doing unnecessary damage to the welfare and
reputation of the group. According to the social identity perspective, groups are engaged in a struggle
for status and positive regard. The outcomes of this contest have implications for self-esteem and in
some cases may also have implications in terms of material rewards. There is an implicit rule, then,
that group members conspire to present the best possible face to the world (see Reicher & Levine,
1994, for a broader discussion of strategic behaviour in groups). This expectation is made explicit by
phrases such as ‘Keep it in-house’, and ‘Don’t air your dirty laundry in public’. Indeed, Scott (1965)
was so convinced of this that he embedded it in his definition of the loyal group member as ‘…a
devoted member of the group, never criticizing it to outsiders, and working hard to get it ahead of other
groups’ (p. 24, italics not in original). Given this, it is reasonable to expect that there would be
normative constraints against criticizing the group to outsiders. Even if people do not have underlying
concerns about the motives for such behaviour, they might punish the critic all the same because they
are unnecessarily placing the reputation and standing of the group at risk.

Although there is no previous research directly testing this notion, the broad principle that group
members engage in strategic self-presentation was demonstrated by Klein and Azzi (2001), who
showed that Belgian participants used more positive and less negative traits to describe their national
character when addressing a French (outgroup) audience than when addressing a Belgian (ingroup)
audience. However, there is also evidence that the need for self-presentation can lead to a more
balanced and objective appraisal of the group. Ellemers, Van Dyck, Hinkle, and Jacobs (2000) showed
that low status group members expressed ingroup bias in private, but outgroup bias when they were
told they would have to discuss their responses to other participants in the experiment (including
outgroup members). They argued that, in the public context, low status group members might feel
compelled to act in line with a consensual understanding of the group’s true value.

THE CURRENT STUDIES

In summary, it has often been assumed that ingroup criticisms are best made in-house, but this notion
has not been subjected to empirical scrutiny. Experiment 1 was designed to help fill this gap in the
literature by examining the role of audience in moderating responses to criticism from ingroup and outgroup members. Experiment 2 went further by seeking to identify why ingroup critics might face heightened defensiveness when they make their comments to an outgroup audience. Finally, Experiment 3 examines the role of identification in moderating responses to criticisms directed to an outgroup audience.

**EXPERIMENT 1**

In Experiment 1, participants from Queensland (a state in Australia) read an extract from an interview in which a person criticized Queenslanders. The comments were attributed to either somebody from Queensland or New South Wales (a neighbouring state of Australia) and participants were told the letter was published in either a Queensland or a New South Wales newspaper. Consistent with the intergroup sensitivity effect, we predicted that participants would show more negativity to the comments when they were made by an outgroup as compared to an ingroup member, and they would agree with the comments less when they were attributed to an outgroup member. We further predicted that audience would moderate the intergroup sensitivity effect. Specifically, when ingroup criticisms were published in an outgroup newspaper, we predicted that participants would show more negativity to the comments and would agree with the comments less than when the criticisms were published in an ingroup newspaper, thus attenuating or eliminating the intergroup sensitivity effect. It was not expected that choice of audience would influence responses to the outgroup critic.

**Method**

*Pilot Study*

A pilot study was conducted to identify credible criticisms for use in the main study. Twenty psychology students living in Queensland were asked to identify three negative characteristics about Queenslanders. The four most frequently named characteristics were narrow-mindedness (10 times), being conservative (three times), being intolerant of other cultures (eight times) and being obnoxious (three times). These characteristics were then incorporated into the experimental scripts.

*Main Study: Participants and Design*

Only Queenslanders were eligible for the study. When asked on a preliminary questionnaire to nominate their ‘primary state identity’, 13 of the original sample recorded a state other than Queensland. These participants were excluded from the study, leaving 116 participants for analysis. The participants consisted of 84 females and 32 males ($M = 20.13$ years) who were randomly assigned to the four conditions of a $2$ (ingroup or outgroup speaker) $\times$ $2$ (ingroup or outgroup audience) between groups design. Participants received course credit in exchange for their time.

*Procedure and Materials*

Participants were told that the study was about judging personalities on the basis of minimal information. They were told that they would be reading an extract of an interview with somebody
talking about Queensland. Participants were instructed to read through these scripts carefully and to fill out the attached questionnaire based on their impressions of the speaker and their comments.

**Manipulations** Depending on condition, participants were told that the comments in the scripts were published in either the Courier Mail (ingroup audience) or the Sydney Morning Herald (outgroup audience). These newspapers were selected because they were the premier broadsheets of Queensland and New South Wales respectively. The participants were told that these statements appeared in the newspaper as a way of marking Australia’s 100th anniversary of Federation; that is, the anniversary of when the states of Australia formally joined together as a nation. Participants were told that the newspaper was trying to gauge, 100 years after Federation, the way Australians feel about people from various states. To do this they had contacted a number of Australians and asked them their opinion of the states. The newspaper had then published extracts from these interviews in a special Federation Anniversary edition of the newspaper. The extract that participants were given read as follows:

> When I think of Queenslanders I think of them as being a bit narrow-minded when it comes to political and environmental issues. They’re resistant to change, and sometimes that can mean being intolerant of other cultures...particularly Asians. Also, I think that they can tend to be pretty obnoxious...especially when drinking.

The scripts were presented with the banner head for the newspaper printed on top of the page. Beneath the printed extract, participants could read the name of the speaker (S. Walker), their place of residence (Brisbane or Sydney, depending on condition), and the state in which they lived. Speakers were identified as being either from Queensland (ingroup critic) or from New South Wales (outgroup critic). When the speaker was a Queenslander, more self-inclusive language was adopted in order to accentuate their ingroup status. For instance, where an outgroup critic said ‘they can tend to be pretty obnoxious’, an ingroup critic said ‘we can tend to be pretty obnoxious’. 

**Measures** After reading the criticism, participants rated the extent to which they felt negative to the criticisms and the extent to which they agreed with the criticisms. These measures—adapted from Hornsey et al. (2002)—were completed using 7-point scales (1 = not at all, 7 = very much). Negativity was measured by asking the extent to which the comments were threatening, disappointing, irritating, offensive, insulting, judgmental, hypocritical, and arrogant (α = 0.88). Participants then rated the extent to which they agreed with the speaker’s comments and the extent to which they thought the speaker’s comments were valid (agreement; r = 0.66). Principal components analysis with varimax rotation confirmed that the items loaded on two factors corresponding to the negativity (eigenvalue = 4.43, 44.34% of variance explained) and agreement (eigenvalue = 2.37, 23.71% of variance explained) measures described above.

**Results**

A series of 2 (Speaker Group: ingroup or outgroup speaker) × 2 (Audience: ingroup or outgroup) ANOVAs revealed main effects of speaker group on negativity, $F(1, 112) = 15.00, p < 0.001, \eta^2 = 0.12$, and agreement, $F(1, 112) = 18.32, p < 0.001, \eta^2 = 0.14$. Consistent with the intergroup sensitivity effect, participants felt more negatively toward criticisms stemming from an outgroup member ($M = 4.95, SD = 1.33$) than from an ingroup member ($M = 3.98, SD = 1.35$), and were less likely to agree with criticisms stemming from an outgroup member ($M = 2.70, SD = 1.57$) than from an ingroup member ($M = 3.90, SD = 1.44$).

On negativity, however, this main effect was qualified by a significant interaction between speaker group membership and audience, $F(1, 112) = 4.07, p = 0.046, \eta^2 = 0.04$ (see Figure 1). Analysis of
simple main effects showed that, when the critic was an ingroup member, participants were more negative toward the comments when there was an outgroup audience ($M = 4.36, SD = 1.19$) than when there was an ingroup audience ($M = 3.64, SD = 1.40$), $F(1, 112) = 4.38, p = 0.039$, an effect that is consistent with predictions. In contrast, when the critic was an outgroup member, there was no significant difference between the level of negativity when there was an outgroup audience ($M = 4.82, SD = 1.24$) and when there was an ingroup audience ($M = 5.09, SD = 1.42$), $F(1, 112) = 0.59, p = 0.44$. One consequence of this is that, although the intergroup sensitivity effect emerged when comments were directed toward an ingroup audience, $F(1, 112) = 17.95, p < 0.001$, there was no difference between ingroup and outgroup critics when comments were made to an outgroup audience, $F(1, 112) = 1.67, p = 0.20$. For ratings of agreement, audience had no effect either independently or as part of an interaction with speaker group.

**Discussion**

Experiment 1 confirmed that, when responding to criticisms of their group, people do not focus just on the content of the criticism. Group members also focus on who is making the criticism, and on the context in which the criticisms are made. Consistent with the intergroup sensitivity effect, participants were more negative to criticisms when they stemmed from an outgroup than an ingroup member. They also agreed with the comments less when they were attributed to an outgroup member than when the very same comments were made by an ingroup member.

More central to the current paper is the fact that choice of audience proved to be influential in determining how people responded to the ingroup critic. As predicted, negativity was higher when the ingroup critic made their comments to outsiders than when they kept their comments in-house, and the effect was large enough to eliminate the intergroup sensitivity effect in the outgroup audience condition. As argued earlier, there are two possible reasons why defensiveness might be heightened when ingroup members criticize their group to outsiders. First, comments that are made to outsiders have a greater potential to damage the reputation and welfare of the group than comments that are made in-house. Second, when insiders take their critical comments to an outgroup audience, group members might question the extent to which the criticisms were made for constructive reasons. The primary aim of Experiment 2 was to examine these possibilities.
Interestingly, although audience influenced ratings of negativity to ingroup criticisms, it had no effect on the extent to which people agreed with the comments. This is not a trivial deviation from the hypotheses. If the critic’s primary objective is to be listened to (as opposed to being liked or maintaining harmony) then the current data suggest that their choice of audience is not relevant after all. One possible reason for the inconsistent effects of audience on the dependent measures is that the power relations between the groups used in Experiment 1 are reasonably equal and benign. It is interesting to speculate whether the same results would emerge if the groups were nested within an unequal power hierarchy. According to social identity theory, evaluations of ingroup and outgroup members differ depending on whether group members perceive alternatives to the existing social order (e.g., Caddick, 1982; Tajfel, 1982; Turner & Brown, 1978). Where power differentials are built on an illegitimate basis, the power hierarchy becomes unstable, inciting attempts to change the status quo. The result is that low power groups might engage in more ethnocentric behaviour as a way of challenging their low power position. When part of an illegitimate power hierarchy, high power groups might also engage in heightened ethnocentrism as a way of justifying or maintaining their high power position (Hornsey, Spears, Cremers, & Hogg, 2003).

When applying these insights to the current research question, it could be that the effects of audience might be more pronounced when groups differ in the amount of power they have. A group with low power within society (particularly for whom the power differential is seen to be illegitimate) might be particularly sensitive to critical comments taken outside the group, because such comments might be used as ‘ammunition’ to help justify the dominant group’s power over them. Even groups with illegitimately high power might show heightened sensitivity to the choice of audience as they respond to the threat inherent in the unstable power hierarchy. In contrast, groups with legitimately high or low power might feel less sensitive about the choice of audience because their groups have less need to engage in strategic self-presentation. A secondary aim of Experiment 2 was to address this question by manipulating not just the audience to whom criticisms are made, but also the power relationship between the groups and the legitimacy of that power relationship.

**EXPERIMENT 2**

Because the focus of the current research questions is firmly on ingroup critics—and because audience had no effect on responses to outgroup critics in Experiment 1—we did not include outgroup critics in the design for Experiment 2. Social science students read a criticism of the social sciences that was attributed to a student representative of the social sciences. Furthermore, they were told that this criticism was made at a closed faculty conference of either social science members (ingroup audience) or maths-science members (outgroup audience).

After reading the scripts, participants rated their negativity toward the comments and the extent to which they agreed with the comments. In addition to these measures we included a number of other scales designed to assess different dimensions of defensiveness. Participants evaluated the critic on a number of personality traits and rated the extent to which they felt the critic had made an appropriate choice of audience. Furthermore, in order to identify the psychological underpinnings of the predicted effect, we included measures of the extent to which the speaker was doing damage to the group, and measures of the extent to which the comments were seen to be motivated by constructive reasons.

Consistent with Experiment 1, it was predicted that participants would express higher levels of negativity, would rate the speaker less favourably on personality traits, and would rate the choice of audience as less appropriate when the comments were made to an outgroup as opposed to an ingroup audience. In line with our arguments earlier, it was further predicted that the effects of audience would
be mediated by ratings of damage to the group and/or by attributions of constructiveness. Because audience had no significant effects on ratings of agreement in Experiment 1, we did not make any firm predictions with regard to how audience might affect participants’ agreement with the comments. Experiment 2 also differed from Experiment 1 in that we manipulated the power relationship between the groups and the legitimacy of that power differential. Working from the assumption that groups nested within an illegitimate power hierarchy have greater need to engage in strategic self-presentation than those nested within a legitimate power hierarchy, we tentatively predicted that the effects of audience described above would be more pronounced in the low as compared to high legitimacy conditions.

Method

Participants and Design

In Experiment 2 ‘social science’ was the ingroup and ‘maths-science’ was the outgroup. Participants were 179 social-science students (136 female, 43 male), aged 17–52 years ($M = 19.75$, $SD = 4.49$), who were randomly allocated to the eight cells of a Power (high or low) × Legitimacy (high or low) × Audience (ingroup or outgroup) between groups design. Participants received either course credit or money ($10) as compensation for their time.

Procedure

Manipulation of Power and Legitimacy Power and legitimacy were manipulated using the technique described by Hornsey et al. (2003). In an initial ‘comprehension task’, participants received information about the university senate. They were told that the senate is divided into two faculty areas: the maths sciences (e.g. biology, engineering) and the social sciences (e.g. arts, psychology). They were told that one purpose of the senate is to make decisions about the division of funding to these faculty areas. In the high power condition, participants were told that eight of the ten senate members represented the social sciences and two represented maths-sciences. In the low power condition this was reversed. Legitimacy was manipulated by comparing the ratio of maths science and social science representatives in the senate with the ratio of maths science and social science students and staff at the University of Queensland (UQ). In the high legitimacy condition, participants were told that the ratio of social science to maths science students and staff at UQ was the same as the ratio within the senate. In contrast, the participants in the low legitimacy condition were informed that the ratio of social science to maths science students and staff at UQ was the reverse of the representation in the senate.

Manipulation Checks Participants were given a comprehension test associated with the first questionnaire, which included questions used as manipulation checks for power and legitimacy. To check whether participants understood the instructions, we asked them to record how many of the University Senate were social science representatives and how many were maths science representatives. To check whether participants internalized the manipulation of power, participants were asked ‘How powerful do you see the maths science faculty to be?’ and ‘How powerful do you see the social science faculty to be?’ (1 = not at all, 9 = very much). To check whether participants internalized the manipulation of legitimacy, they were asked ‘To what extent do you feel the division of maths science to social science representatives in the University Senate is fair?’ (1 = not at all, 9 = very much).
Manipulation of Audience Audiences group membership was manipulated in a second questionnaire where participants were given a criticism of the social sciences to read. They first received background information on the critic, which included their name and their position. The critic’s position was stated as a student representative of the social sciences faculty area. Participants were also given background information on the critic’s comments, indicating that it was part of a speech taken from a conference of either social science (ingroup audience), or maths-science (outgroup audience) faculty members. In an effort to make the manipulation of audience as clean as possible, they were told that the conference was a closed meeting with only members of each relevant faculty area in attendance. Participants then read the criticism:

As social science students we don’t do enough to make ourselves relevant to society; we struggle to prove our worth to others. I also think that we are slow to make strategic links with industry and consequently have become too reliant on government funding. Social science students appear lost in an ivory tower, isolated from the real world.

Audience Manipulation Check After reading the information about the critic, participants completed a questionnaire involving evaluations of the critic and their comments, as well as three questions used as manipulation checks for audience: ‘Of which faculty is the speaker a student representative?’, ‘Where were the speaker’s comments made?’, and ‘Who was the audience for the speaker’s comments?’. 

Dependent Measures Subsequent items—adapted from Hornsey et al. (2004)—were assessed using 7-point scales (1 = not at all, 7 = very much). The extent to which the speaker’s comments were seen to be doing damage to the group was assessed by asking: ‘To what extent do you feel that the speaker is bad for the welfare of social science students?’ and ‘To what extent do you feel that the speaker has done damage to the reputation of social science students?’ (damage; \( \alpha = 0.82 \)). Three questions measured the extent to which the speaker's comments were attributed to constructive motives (constructiveness): ‘To what extent do you think the comments were intended to be constructive?’, ‘To what extent do you think the comments were intended to be destructive?’, and ‘To what extent do you think the comments were made in the group’s best interests?’ (\( \alpha = 0.80 \)). Personality evaluations were measured by asking participants the extent to which they believed the speaker to be: intelligent, friendly, open-minded, likable, respected, interesting, and nice (\( \alpha = 0.88 \)). Negativity items were the same as those used in Experiment 1 (\( \alpha = 0.87 \)).

One concern we had with Experiment 1 is that the measure of agreement reflected a broad, subjective sense of correspondence with the argument (e.g. ‘To what extent do you agree with the comments?’), ‘To what extent do you think the comments are valid?’) and did not weight heavily enough indices of whether participants believed the specific comments raised by the critic were objectively true. In response to this, we replaced the two-item measure of agreement used in Experiment 1 with a more sophisticated four-item measure. This included one item designed to capture globally the extent to which participants believed the comments to be based in fact (‘To what extent do you believe the comments to be true?’) and three items designed to capture agreement with each of the three specific criticisms raised by the critic (‘To what extent do you feel social science students need to do more to make themselves relevant to society?’, ‘To what extent do you feel social science students need to work to make more strategic links with industry?’ and ‘To what extent do you feel social science students need to work to dispel society’s beliefs that they are isolated from the real world?’). Together, these four items formed an internally consistent scale (agreement; \( \alpha = 0.85 \)). Finally, three questions assessed the appropriateness of the audience: ‘To what extent do you think the speaker chose the right audience to make these comments?’ ‘To what extent do you think the speaker chose an appropriate forum to make
their comments?’ and ‘To what extent do you think the speaker should have chosen a different audience to make their comments?’ (The last question was reverse scored; $\alpha = 0.87$).

Results

Preliminary Analyses

To test whether the dependent measures represent discrete constructs, a factor analysis was conducted using principal components extraction with varimax rotation. Initial analyses showed that two items in the negativity scale (i.e. ratings of the extent to which the comments were threatening and arrogant) loaded on more than one factor, and so these items were discarded (results on the negativity measure were the same regardless of whether these two items were included or deleted). In subsequent analyses, a five-factor solution emerged with the first factor comprising the personality evaluations (eigenvalue = 8.64, 34.58% of variance explained), the second factor comprising the negativity items (eigenvalue = 2.65, 10.61% of variance), the third factor comprising the agreement items (eigenvalue = 2.27, 9.08% of variance), the fourth factor comprising the damage and constructiveness items (eigenvalue = 2.01, 8.04% of variance), and the fifth factor comprising the appropriateness of the audience items (eigenvalue = 1.27, 5.09% of variance). Although the damage and constructiveness measures are highly correlated ($r = -0.60$), it would be difficult to justify combining them into a single scale for face validity reasons. Consequently, we chose to analyse the six constructs separately.

Manipulation Checks

Two participants were removed because they incorrectly cited the ratio of social science and maths-science representatives in the senate, and three further participants were removed because they failed the audience manipulation checks. To check on the manipulation of power, a difference score was calculated by subtracting power ratings for the maths sciences faculty area from power ratings for the social sciences faculty area. A 2 (Power: high or low) $\times$ 2 (Legitimacy: high or low) $\times$ 2 (Audience: ingroup or outgroup) between groups ANOVA confirmed that the manipulation of power was successful, $F(1, 165) = 408.91, p < 0.001, \eta^2 = 0.71$, such that in the high power condition our social science participants saw their group to be more powerful than the maths-sciences ($M = 3.31, SD = 2.94$), whereas in the low power condition participants saw their group to be less powerful than the maths-sciences ($M = -4.32, SD = 2.22$). Contrary to expectations, a much weaker main effect of legitimacy also emerged, $F(1, 165) = 12.50, p < 0.001, \eta^2 = 0.07$, such that participants in the high legitimacy condition ($M = 5.58, SD = 2.41$) saw their power position to be more fair than did those in the low legitimacy condition ($M = -0.94, SD = 4.59$).

Analysis on the legitimacy manipulation check revealed a large main effect of legitimacy, $F(1, 166) = 110.81, p < 0.001, \eta^2 = 0.40$, such that participants in the high legitimacy condition ($M = 5.58, SD = 2.41$) saw their power position to be more fair than did those in the low legitimacy condition ($M = 2.65, SD = 1.43$). A weaker main effect of power also emerged, $F(1, 166) = 14.77, p < 0.001, \eta^2 = 0.08$, such that participants in the high power condition ($M = 4.54, SD = 2.47$) perceived their power position to be more fair than did those in the low power condition ($M = 3.62, SD = 2.36$). This slight cross-over between the manipulation of power and legitimacy was found also by Hornsey et al. (2003) and presumably reflects a motivation for high power groups to rationalize their positions as fair (see Chen & Tyler, 2001, for a discussion of legitimizing beliefs). Nevertheless,
it is clear that the manipulations of power and legitimacy had dramatic effects on perceptions of relative power and fairness of the power differential respectively, indicating that the manipulations were internalized by participants.

Main Analyses

To examine whether audience influenced responses to group-directed criticism, a series of ANOVAs was conducted on negativity, personality evaluations, appropriateness, and agreement (see Table 1 for summary of means and standard deviations). Consistent with expectations, significant main effects of audience were found for ratings of personality evaluations, $F(1, 166) = 4.50, p = 0.035, \eta^2 = 0.03$, and appropriateness, $F(1, 166) = 46.82, p < 0.001, \eta^2 = 0.22$. Critics who kept their comments in-house were evaluated more positively on personality traits and were seen to have made a more appropriate choice of audience than were those who made their comments to an outgroup audience. On ratings of agreement, a marginal effect of audience emerged, such that participants expressed stronger agreement when the comments were made to an outgroup audience than when they were made to an ingroup audience, $F(1, 166) = 2.94, p = 0.088, \eta^2 = 0.02$. No effects of power or legitimacy emerged on ratings of agreement, personality evaluations, or appropriateness.

Analysis of the negativity ratings revealed a significant interaction between legitimacy and audience, $F(1, 166) = 5.06, p = 0.026, \eta^2 = 0.03$. Negativity ratings in the low legitimacy condition were equivalent regardless of whether criticisms were made to an ingroup ($M = 4.16, SD = 1.26$) or an outgroup ($M = 3.85, SD = 1.31$) audience, $F(1, 166) = 1.16, p = 0.28$. However, in the high legitimacy condition, comments directed toward the outgroup audience ($M = 4.70, SD = 1.48$) aroused more negativity than did comments directed toward the ingroup ($M = 4.09, SD = 1.23$), $F(1, 166) = 4.38, p = 0.038$.

The Mediating Role of Constructiveness and Damage

To examine the psychological underpinnings of the effects of audience on the dependent measures, ANOVAs were conducted on damage and constructiveness. A main effect of audience emerged on
ratings of damage, \( F(1, 166) = 3.95, p = 0.049, \eta^2 = 0.02 \), such that ingroup critics who made their comments to an outgroup audience were seen to be doing more damage to the group than were those who made their comments to an ingroup audience. In terms of constructiveness of the comments, however, only an interaction was found between power and audience group membership, \( F(1, 166) = 4.35, p = 0.039, \eta^2 = 0.03 \). Simple main effects revealed that, when the speaker was from a low power group, perceptions of constructiveness did not differ between the ingroup (\( M = 4.24, SD = 1.69 \)) and outgroup (\( M = 4.43, SD = 1.36 \)) audience conditions, \( F(1, 166) = 0.34, p = 0.56 \). However, when the speaker was from a high power group, comments made to the outgroup (\( M = 4.04, SD = 1.23 \)) were viewed as less constructive than were comments made to the ingroup (\( M = 4.73, SD = 1.31 \)), \( F(1, 166) = 5.59, p = 0.019 \).

In summary, ingroup critics were evaluated less positively and their choice of audience was seen to be less appropriate when they spoke to an outgroup audience than when they kept their comments in-house. Because there was no main effect of audience on constructiveness, the mediating role of this variable was not pursued further. To examine the mediating role of damage, a series of regressions was conducted. When audience (dummy coded) and damage were entered in the model simultaneously, damage was a significant predictor of both personality evaluations (\( \beta = -0.56, p < 0.001 \)) and appropriateness (\( \beta = -0.27, p < 0.001 \)). After including damage in the regression equation, a significant effect of audience on personality evaluations (\( \beta = -0.16, p = 0.035 \)) was rendered non-significant (\( \beta = -0.08, p = 0.23 \)). The reduction in variance caused by entering damage as a mediator was marginally significant according to the Sobel test (\( z = -1.93, p = 0.054 \)). Damage was less successful, however, in explaining the effects of audience on appropriateness. After including damage in the regression equation, the significant effect of audience on appropriateness (\( \beta = -0.47, p = 0.001 \)) was reduced but remained highly significant (\( \beta = -0.43, p = 0.001 \)). The reduction in variance caused by entering damage as a mediator was marginally significant (\( z = -1.78, p = 0.076 \)).

**Discussion**

Participants felt that the choice of audience was less appropriate when the comments were made outside the group, and rated the speaker less positively on personality traits when they made their comments to outsiders than when they kept their comments in-house. Furthermore, Experiment 2 was able to provide insight into why the choice of audience might have an impact on reactions to criticism. Compared to criticisms that were kept in-house, comments that were taken to outsiders were seen to be doing more damage to the welfare and reputation of the group, and it was this tendency that underpinned the effects of audience on personality evaluations. There was no evidence that the effects of audience were driven by attributions; indeed, choice of audience had no effect at all on whether the speaker was attributed constructive motives for their criticisms. In general, then, participants seem to be characterizing the speaker who had taken their comments to outsiders as well intentioned but potentially threatening to the integrity of the group, and for this reason the speaker was downgraded on personality dimensions.

It should be noted that the downgrading of the speaker in the outgroup audience condition, although statistically significant, was not dramatic in size. Furthermore, choice of audience did not influence ratings of agreement. In summary, although participants felt strongly that critics who took their comments outside the group had behaved inappropriately in their choice of audience, this had only modest consequences for them in terms of how positively they were evaluated, and had no significant consequences in terms of whether people agreed with the comments.

The mixed effects of audience—and in particular the failure of audience to influence ratings of agreement—resemble the pattern of results found in Experiment 1. We speculated that one reason
why audience might have had limited effects in Experiment 1 is that we used groups with relatively equal power. Part of the rationale for Experiment 2 was to test the notion that groups with low power—and particularly groups for whom the power differential is illegitimate—might be more negative than other groups to having criticisms taken to outsiders. But there was no evidence for this in the current study. Comments that were taken to outsiders aroused more negativity than did comments kept in-house, but only for groups within legitimate power hierarchies. Critics who spoke to outsiders were attributed less constructive motives for their comments than were those who kept their comments in-house, but only for groups with high power. If anything, it was the groups that we assumed had the most to lose—those with low power and those nested in illegitimate power differentials—who were least defensive in the outgroup audience condition. It is not immediately clear why this should be the case. One possibility is that members of groups in the low power conditions and the illegitimate power conditions already felt as though their group was in a hostile intergroup context, and so the failings of the group were already an ‘open secret’. As such, presentation of these failings to an audience of outgroup members might do little to harm their image, and in fact might have benefits in terms of projecting an impression of self-assurance, self-confidence, and commitment to improvement. Whatever the specific mechanisms that are in operation here, there was no evidence that the effects of audience could be accentuated by focusing on groups with low power or with illegitimate power.

In summary, the results of Experiments 1 and 2 conformed to many of our expectations. Criticisms that are made to an outgroup audience arouse greater negativity (Experiment 1) and are seen to be more inappropriate (Experiment 2) than are comments kept in-house. Furthermore, critics who take their comments to outsiders are rated more negatively on personality evaluations than are those who seek an ingroup audience (Experiment 2). Presumably, however, a person rarely makes critical comments about a group to be liked or to protect group members’ feelings, so the negativity directed towards the speaker in the outgroup audience condition might not be of great consequence. The ultimate goal of a critic is to promote change, and so the critical question becomes: How does audience affect the extent to which group members agree with the criticisms? Examination of the data in Experiment 2 provides an unexpected answer to this question. Despite the fact that the speaker was seen to have made an inappropriate choice when they took their comments to outsiders, despite the fact that comments taken to outsiders were seen to be doing more damage to the group, and despite the fact that the speaker was liked less when they took their comments outside the group, these comments were marginally more successful in winning over agreement than were comments kept in-house. Why should the effects of audience on agreement defy the trend established on the other indices of defensiveness?

Commitment To Reform Versus Denial

In answering this question, one should consider the choices that are faced by a person who has just witnessed their group being criticized to outsiders. One option is to discredit the critic and to go into denial about any truth that might lie within the comments. There are three reasons, however, to believe that this option might not be satisfactory. First, the comments were based on pilot studies conducted on group members themselves, and so presumably carry with them at least a kernel of truth. Second, participants perceive the critic to be motivated by relatively constructive motives even when they make their comments to an outgroup audience. Previous research on group-directed criticism has revealed that suspicion about motive is a major avenue through which people discredit criticisms (e.g. Hornsey & Imani, 2004). Given that trust towards the ingroup critic seems robust in the current studies, participants are unable to rely on this rationalization for why the criticisms should be dismissed. Third,
and perhaps most importantly, denial does nothing to stem the damage the criticisms have caused to the reputation and welfare of the group (see also Ellemers et al., 2000). If anything, it might leave the group vulnerable to accusations of being defensive and immature.

An alternative option to denial is to embrace the truth within the comments and to work to correct them. The great advantage of this strategy is that it addresses the problem at its root cause. It provides a constructive, long-term response to the damage that might have been caused by the criticism, and helps protect the group against damage that might be caused by criticism in the future. Not only does it help protect the reputation of the group, it also becomes an avenue through which the group’s culture can be reformed, the positive consequences of which presumably flow on to all group members. In the face of legitimate criticism, acknowledging the truth within the comments and committing the group to reform would seem to be the ideal response because it strengthens the group and protects it from further attack. Although criticisms made to outgroup audiences might arouse a degree of emotional negativity, the heightened damage such criticism causes might create a catalyst for action that would not necessarily occur to the same degree if criticisms were kept in-house. Commitment to reform, however, would come with costs; it requires genuine effort, and the positive effects of reform are unlikely to ameliorate the damage that is caused in the short-term.

If this reasoning is correct, one might expect that the effects of having one’s group criticized to an outgroup audience might differ depending on the extent to which the group member identifies with their group. Previous research has shown that high and low identifiers often behave differently when the group is facing a threat to its identity (see Branscombe, Ellemers, Spears, & Doosje, 1999, for a review). For example, when forced to make an unflattering status comparison with an outgroup, high identifiers are more likely than low identifiers to describe themselves in line with the group stereotype (Spears, Doosje, & Ellemers, 1997), more likely to rate the group as homogenous (Doosje, Ellemers, & Spears, 1995), and less likely to seek escape from the group in order to join the higher status outgroup (Ellemers, Spears, & Doosje, 1997). In short, these studies have found that low identifiers are more likely to opt for individual mobility and other individual-level identity management strategies in the face of threat, whereas high identifiers are more likely to employ group-level strategies to collectively respond to threat.

In many ways, the outgroup audience condition of the current studies resembles an identity threat condition. By exposing the group’s shortcomings to an outgroup audience, group members are forced to reflect in an unflattering way on the strengths and weaknesses of their group relative to the outgroup. Extrapolating from previous research on identity threat, high identifiers would be more likely than other group members to respond to this threat by engaging in collective efforts to improve the group. Specifically, when criticisms are made to an outgroup audience, one might expect that high identifiers would be more likely than low identifiers to embrace the critic and their comments, whereas low identifiers might be more likely to engage in intrapsychic or interpersonal strategies such as rejecting the comments as untrue or distancing the self from the critic. Experiment 3 was designed to test this prediction.

EXPERIMENT 3

Social science students were led to believe that their group had been criticized to either an ingroup or an outgroup audience, using the same paradigm used in Experiment 2. Participants’ identification with the social science group was measured prior to exposure to the criticism. In line with Experiment 2, it was predicted that the choice of audience would be seen to be more appropriate when the comments
were delivered to an ingroup as compared to an outgroup audience. We also predicted that critics would be evaluated more positively when the comments were made to an ingroup audience. In line with Experiments 1 and 2, we did not expect that the group membership of the audience would have an independent effect on ratings of agreement.

We did, however, predict that how participants responded to being criticized in front of ingroup and outgroup members might differ as a function of participants’ identification levels. We expected that those with only moderate identification with the group would respond to the identity threat associated with being criticized to the outgroup by denying the truth within the comments (an intrapsychic strategy) or by distancing the self from the critic (an interpersonal strategy). Specifically, it was predicted that levels of agreement with the criticisms and friendly intentions toward the critic would be lower in the outgroup audience condition than in the ingroup audience condition. For high identifiers, we expected this tendency to be attenuated or even reversed as they respond in a collective fashion to the identity threat represented by having criticisms made to the outgroup audience. Thus, for high identifiers, we expected that levels of agreement and friendly intentions would be just as strong—if not stronger—in the outgroup audience condition as in the ingroup audience condition.

Method

Participants and Design

Fifty nine first year social science students studying social sciences at UQ participated in the experiment for course credit. Participants consisted of 42 females and 17 males (\(M = 21.17\) years, \(SD = 6.14\)) who were randomly assigned to either an ingroup or an outgroup audience condition.

Procedure and Measures

After recording their biographical details, participants completed a three-item identification scale. On 9-point scales (1 = not at all, 9 = very much) participants recorded the extent to which they agreed with the following statements: ‘I identify as a social science student’, ‘I have a lot in common with social science students’, and ‘Being a social science student is an important part of my self-image’ (\(\alpha = 0.83\)). Because it is possible that feelings toward the ingroup speaker might differ as a function of identification—indeed, of whether the speaker criticized the group—we felt it would be beneficial to gain a baseline measure of personality evaluations and friendly intentions before the criticisms were revealed. Participants were told that they would soon be given an extract from a speech from a student representative of the social sciences (T. Robertson) that was made at either a social science or a maths science conference (depending on condition). Participants then rated the speaker using the same 7-item scale of personality evaluations used in Experiment 2 (\(Time 1\) personality evaluations; \(\alpha = 0.91\)). These items—and all subsequent items—were responded to using 7-point scales (1 = not at all, 7 = very much). In addition to the personality evaluations used in Experiment 2, we also included two items that were designed to capture more specifically a sense of engagement and attraction to the speaker. Participants rated the extent to which they agreed or disagreed with the following statements: ‘I would be interested in getting to know this person’ and ‘I would be interested in being this person’s friend’ (\(Time 1\) friendly intentions; \(r = 0.85\)).

Participants were then exposed to the same criticism of social science students used in Experiment 2. After reading the comments, participants filled out the personality evaluations and friendly intentions scales again (\(Time 2\) personality evaluations; \(\alpha = 0.91\); \(Time 2\) friendly intentions; \(r = 0.85\)).
Results

Hierarchical multiple regression was performed, with identification and audience entered in the first step (audience dummy coded such that ingroup audience = 0, outgroup audience = 1) and the interaction term between identification and audience entered in the second step (in line with recommendations by Aiken & West, 1991, continuous variables were centred). When performing regressions on ratings of personality evaluations and friendly intentions, Time 2 scores were entered as criteria but with Time 1 scores entered as a predictor at the first stage.

Consistent with expectations, an effect of audience emerged on ratings of appropriateness ($\beta = -0.53$, $p < 0.001$), such that the choice of audience was seen to be more appropriate when criticisms were delivered to an ingroup ($M = 4.76$, $SD = 1.44$) as opposed to an outgroup audience ($M = 2.83$, $SD = 1.63$). No other main effects emerged. There was, however, a significant interaction between identification and audience on ratings of agreement ($\beta = 0.49$, $p = 0.008$) and friendly intentions ($\beta = 0.46$, $p = 0.003$). We analysed simple slopes by performing regressions separately at each level of audience. When the criticisms were delivered to an ingroup audience, there was no relationship between identification and agreement ($\beta = -0.16$, $p = 0.42$) or identification and Time 2 friendly intentions ($\beta = -0.22$, $p = 0.21$). When the criticisms were delivered to an outgroup audience, however, a significant positive relationship emerged, such that the more participants identified with their group the more they agreed with the comments ($\beta = 0.49$, $p = 0.006$) and the more they expressed friendly intentions toward the speaker ($\beta = 0.51$, $p = 0.001$).

For reasons of interpretability, we then performed a median split on identification (moderate identifiers: $M = 3.34$; high identifiers, $M = 6.06$) and conducted a $2 \times 2$ between-groups ANOVA on the agreement measure. Consistent with the regression described above, a marginally significant interaction emerged between identification and audience on agreement, $F(1, 55) = 3.86$, $p = 0.054$, $\eta^2 = 0.07$ (see Figure 2). As predicted, moderate identifiers expressed more agreement when the comments were made to an ingroup audience than when they were made to an outgroup audience, $F(1, 55) = 4.47$, $p = 0.039$. In contrast, high identifiers’ agreement levels were not affected by audience, $F(1, 55) = 0.43$, $p = 0.51$.

Finally, a 2 (ingroup v/ outgroup audience) $\times$ 2 (high v/ moderate identifier) $\times$ 2 (Time 1 v/ Time 2 ratings) mixed ANOVA was conducted on the ratings of friendly intentions, revealing a main effect of time, $F(1, 55) = 8.00$, $p = 0.007$, $\eta^2 = 0.13$, and identification, $F(1, 55) = 4.04$, $p = 0.049$, $\eta^2 = 0.07$. Overall, participants had less friendly intentions toward the target after receiving the criticisms ($M = 3.43$, $SD = 1.32$) than before ($M = 3.81$, $SD = 1.29$). Furthermore, high identifiers ($M = 3.68$, $r = 0.92$). Participants also reported how negatively they felt toward the comments ($\text{negativity; } \alpha = 0.88$), how constructive the comments were seen to be ($\text{constructiveness; } \alpha = 0.85$), and how appropriate the comments were seen to be ($\text{appropriateness; } \alpha = 0.94$). All scales used the same items to those used in Experiment 2. Finally, participants responded to the four agreement items used in Experiment 2 as well as a new measure: ‘To what extent do you feel social science students need to act to correct the problems raised by the speaker?’ These five items inter-correlated highly and so were combined into a single scale of agreement ($\alpha = 0.87$). Factor analysis with varimax rotation was performed on all the items measured at Time 2. As expected, a six-factor solution emerged, with factors corresponding to the scales of Time 2 personality evaluations (eigenvalue = 7.73, 29.74% of variance explained), negativity (eigenvalue = 4.45, 17.11% of variance), agreement (eigenvalue = 2.91, 11.18% of variance), appropriateness (eigenvalue = 2.58, 9.92% of variance), constructiveness (eigenvalue = 1.19, 4.59% of variance), and Time 2 friendly intentions (eigenvalue = 1.11, 4.28% of variance).
SD = 1.52) expressed stronger friendly intentions overall than did moderate identifiers (M = 3.17, SD = 1.04). These main effects, however, were qualified by a significant interaction between time, identification, and audience, F(1, 55) = 9.49, p = 0.003, \( \eta^2 = 0.15 \). Follow-up analyses showed no main effects or interactions at Time 1. In contrast, a significant interaction emerged between identification and audience on Time 2 ratings, F(1, 55) = 17.20, p < 0.001, \( \eta^2 = 0.24 \) (see Figure 3). Consistent with the effects on agreement, moderate identifiers expressed stronger friendly intentions when the comments were made to an ingroup audience than when they were made to an outgroup audience, F(1, 55) = 12.30, p = 0.001. High identifiers, in contrast, expressed stronger friendly intentions toward the critic when the comments were made to an outgroup audience than when they were made to an ingroup audience, F(1, 55) = 5.54, p = 0.022.

**Discussion**

As in Experiment 2, comments were seen to have been more appropriate when made to an ingroup audience than when made to an outgroup audience. However, there was limited evidence that the
ingroup critic who made their comments to an outgroup audience was punished for this strategic error. Compared to critics who kept their comments in-house, the critic who chose an outgroup audience for their comments was evaluated just as positively, was attributed equally constructive motives, aroused no more negativity, and was just as successful in winning agreement for their arguments.

Closer analysis, however, revealed that the choice of audience was having an effect on responses to criticism, but that these effects were moderated by participants’ identification levels. People who criticized their group to an outgroup audience won over less agreement and weaker friendly intentions than did those who kept their criticisms in-house, but only among moderate identifiers. High identifiers, in contrast, were less punitive of the person who made their comments to an outgroup audience. They were just as likely to agree with the comments when they were made to an outgroup as opposed to an ingroup audience, and they expressed significantly stronger friendly intentions toward the person who made their comments to an outgroup audience as opposed to an ingroup audience. It should be noted that identification did not moderate the effects on constructiveness, personality evaluations, or negativity. In other words, when participants were interpreting the situation (‘Why is the person saying this?’, ‘What kind of person is this?’, ‘How do I feel about these comments?’) identification did not play a systematic role. But when deciding how to respond to the comments (‘Do I embrace the comments and accept the need for reform or do I deny the comments?’, ‘Do I reject the critic or embrace them?’) then identification became highly influential. In short, when the criticisms were made to an outgroup audience, high identifiers tended to align themselves to the critic and acknowledge the need to address the group’s shortcomings more so than moderate identifiers.

On the basis of a superficial examination of the literature on ingroup deviancy, this finding might seem somewhat surprising. For example, research on the black sheep effect shows that it is high identifiers—rather than low identifiers—who are most likely to show heightened rejection of deviant ingroup members (Branscombe, Wann, Noel, & Coleman, 1993). The argument is that the non-normative behaviour of the deviant ingroup member threatens the reputation and standing of the group in the eyes of relevant outgroups. The response to this is to show heightened hostility towards the deviant ingroup member, an effect that is most pronounced when the identity is important to people’s self-concept. Why, then, in the current study are high identifiers so supportive of the ingroup critic who seeks an outgroup audience, and in so doing does the most to damage the group’s reputation in the eyes of outgroup members?

The simple answer to this is that an ingroup critic is qualitatively different from the incompetent, dislikable, or disloyal group members typically examined in the literature on the black sheep effect. For these ingroup deviants, the threat is exclusively a result of the presence of the unattractive individual, who by association reflects poorly on the wider group. There are some similarities between these types of ingroup deviants and the ingroup critic, but there are also many important differences. Although criticism is inherently aversive, legitimate criticism can also be valuable because it suggests avenues for improvement. Reform and growth often begin with acts of criticism that are embraced and used as catalysts for positive change. Thus, legitimate criticism represents threat and opportunity in equal measures. Furthermore, research on the intergroup sensitivity effect shows that ingroup critics are not subjected to the high levels of defensiveness faced by other critics because people place relatively high levels of trust in their motives (Hornsey & Imani, 2004). When it comes to ingroup criticism, then, it can be argued that the real source of threat is not the critic per se, but rather the weaknesses and shortcomings within the group, upon which the critic is merely shining a light. In other words the threat is not specific to the deviant individual but strikes more generally at the integrity of the group culture. When criticisms are visible to an outgroup audience the effect resembles an identity threat manipulation.

When viewed as identity threat, the effects of identification observed here do not seem so surprising. Indeed, the moderating role of identification seems consistent with other research on
how group members respond to threatened identity (Doosje et al., 1995; Ellemers et al., 1997; Spears et al., 1997). The message from this literature is that high identifiers are more likely to respond to identity threat through collective means, whereas low identifiers are more likely to choose individual identity management strategies. This is consistent with the findings of Experiment 3. When group members face the greatest threat (i.e. when criticisms are communicated to an outgroup audience) high identifiers take a collective strategy; compared to weaker identifiers they are more likely to rally around the critic and acknowledge the need to address the group’s problems. Moderate identifiers, however, are more likely to respond to this situation with strategies that are more intrapsychic (i.e. denying the truth within the comments) or interpersonal (distancing the self from the critic). This finding should be taken in the context of an emerging body of work showing the ambivalence that group members feel toward those within their ranks who stand up and articulate the group’s shortcomings and failings (e.g. Hornsey et al., 2002, 2004). Clearly, ingroup critics unleash a complex array of feelings and perceptions that are not readily understood from the perspective of related literature on ‘black sheep’.

**GENERAL DISCUSSION**

It is widely assumed that criticisms of groups should be made in-house; indeed, some argue that taking group criticism to outsiders is the essence of disloyalty (Scott, 1965). Such prescriptions are embedded in folk wisdom (‘Never air your dirty laundry in public’) but have not, until now, been subjected to empirical attention. Three experiments designed to examine this notion found some evidence that group members do, indeed, feel heightened defensiveness toward those who take criticisms of the group to outsiders.

The strongest evidence for this notion comes on ratings of appropriateness, upon which the manipulation of audience had a large and consistent effect. In Experiments 2 and 3, critics who kept their comments in-house were seen to have made a much more appropriate choice of audience than were those who took their comments to an outgroup audience. There was also some evidence that choice of audience had deeper consequences for how group members felt about the critic and their message. There was evidence that critics aroused more negativity when they made their comments to an outgroup as compared to an ingroup audience (although this effect was limited to Experiment 1 and to the high legitimacy condition of Experiment 2). Furthermore, critics who took their comments to outsiders were rated more negatively on personality evaluations than those who sought an ingroup audience (Experiment 2). Mediational analysis showed that when comments were taken to outsiders, group members perceived that they were doing more damage to the group, and it was this that underpinned the downgrading of the critic in the outgroup audience condition. Although these effects were weaker and less consistent than the effects on appropriateness, the overall picture remains in line with predictions; where audience had an effect, it was such that critics met less defensiveness when keeping their criticisms in-house.

But there is an important caveat to the ‘intuitive’ picture of how audience might affect responses to ingroup critics. Although participants seemed to disapprove of the ingroup critic taking their comments to outsiders, it did not affect the extent to which they internalized the critic’s message. Participants were no less likely to agree with comments when they were made to an outgroup audience, and there was even a marginal tendency in Experiment 2 for participants to express more agreement with comments when they were made to an outgroup audience than when they were kept in-house. Experiment 3 showed that choice of audience does have an effect on responses to criticism, but that this effect is moderated by identification. When criticisms were taken to an outgroup audience,
high identifiers were more likely than other group members to rally behind the critic and acknowledge the need to address the shortcomings of the group. In fact, high identifiers expressed more friendly intentions toward a critic who took their comments to an outgroup audience than to a critic who kept their comments in-house. Weaker identifiers, on the other hand, were more likely to deny that the comments were true and to distance the self from the critic if they took their comments to an outgroup as compared to an ingroup audience. If one interprets ingroup criticism as a form of identity threat, this finding reinforces a welter of research showing that high identifiers tend to respond to identity threat using unified, collective strategies whereas low identifiers tend to respond using more intrapsychic and inter-individual strategies (see Branscombe et al., 1999, for a review).

It is notable that, although participants expressed quite strongly that it is more appropriate to keep criticisms in-house, violations of this norm had relatively weak and inconsistent effects on evaluations of the critic and their comments. Perhaps one factor that led to the muted evaluations of the insider who takes their comments to the outgroup is that they were still attributed relatively constructive motives for their comments. Indeed, ingroup critics were attributed relatively constructive motives throughout and this did not change when they took their comments to an outgroup audience. Overall, then, the picture portrayed by the data is of somebody who is well intentioned but had made a strategic mistake. Indeed, a prevailing theme that runs through the literature on group-directed criticism (Hornsey et al., 2002; Hornsey & Imani, 2004) is how trusting group members are of the motives of people who stand up and criticize their own group. Although such comments must raise anxiety and threaten the short-term harmony of the group, people seem remarkably willing to take such criticisms in the spirit of constructiveness, even when they are made to an outgroup audience.

It should be noted, though, that the critic in Experiments 2 and 3 fulfilled other criteria that would have helped shore up their credentials in the eyes of other members (see Hornsey et al., 2004). First, they used inclusive, identity-embracing language when making the criticisms, implying an on-going commitment to the group. Second, the speaker had demonstrated their commitment to the group by serving as a representative for the social sciences. Even though it could be argued that there would be strong normative pressures for leaders such as this to not criticize the group to outsiders, the fact they were a leader may have benefited them in the sense that it suggests strong commitment to the group, a quality that has previously been found to be crucial in leading to open-minded responses to ingroup critics (Hornsey et al., 2004). If the speaker’s commitment to the group was more questionable, it is possible that they would not receive as positive a reception. This may be a reason why the effect of audience on negativity emerged in Experiment 1—in which the critic was a rank-and-file member—and not in Experiments 2 and 3, where the critic played a leadership role.

Furthermore, it is worthwhile speculating on how the intergroup context might influence the effects of audience. In both the experiments conducted here the intergroup context was characterized by rivalry but not by open hostility. It could be that, if the groups had been engaged in extreme conflict (e.g. war), members might have been more sensitive to the critic’s choice of audience. In other words, they might have been less forgiving of an insider who takes criticisms to a despised enemy than to those who take their criticisms to a friendly rival.

Another point that should be noted is that, in both cases, the negative qualities identified by the ingroup speaker are widely recognized outside the group. For example, the criticisms of Queenslanders used in Experiment 1 are consistent with a widely held national stereotype of Queenslanders as conservative and uncultured. Similarly, the criticisms used in Experiments 2 and 3 are consistent with a commonly held accusation that the social sciences struggle for relevance and are out of touch with the real world. In these cases, it is possible that the choice of audience does not have a large emotional impact for group members, because they reason that the outgroup audience already holds these views. Indeed, as we argued earlier, acknowledging these shortcomings to an outgroup audience might project an image that the group is mature, self-assured, and actively seeking positive change. On some
occasions, however, criticisms might expose shortcomings of the group that had not been previously recognized by the outgroup; for example, where a member of an organization publicly announces cases of corruption within the organization, or where activists alert the world to human rights abuses in their country. On these occasions, the critical comments might be seen to be particularly damaging to the reputation of the group because it is revealing an unflattering secret. This, in turn, might lead to more intense emotional reactions on behalf of group members.

In summary, the experimental evidence reported here presents a nuanced picture of the role of audience in affecting responses to criticism. Although taking criticisms to outsiders is seen to be inappropriate and can cause internal ructions, it can be reasonably effective in winning support and acting as a catalyst for change, at least among those who identify strongly with the group. Thus, although the studies presented here generally support the folk wisdom that it is best to keep criticisms in-house, the data also present theoretically interesting caveats to this ‘intuitive’ picture of how audience might affect responses to group-directed criticism.

ACKNOWLEDGEMENTS

This research was supported by an Australian Research Council-Discovery grant to the first author.

REFERENCES


