Forewarnings of Influence Appeals: Inducing Resistance and Acceptance

Jeffrey M. Quinn
Wendy Wood
Texas A&M University

According to conventional wisdom, “forewarned is forearmed.” That is, warning of an impending request allows people to prepare for it and ultimately to resist it. For instance, advance knowledge that a telemarketer is about to call and deliver an unwanted sales pitch or that a friend is about to ask a burdensome favor should allow the target of such appeals to mount a successful defense. The idea that warnings generate resistance also is evident in reviews of persuasion research, which typically discuss forewarning effects along with other resistance techniques (e.g., Eagly & Chaiken, 1993). The assumption that warnings yield resistance can also explain a common practice in psychology experiments on attitude change. Experimenters often avoid warning participants of an impending persuasive communication, presumably to maximize participants’ susceptibility to persuasion (Papageorgis, 1967, 1968).

In the present chapter, we consider whether warnings do in fact yield resistance to impending appeals—whether people who are forewarned are better able to resist temptation and stand firm in their beliefs. We review the experimental research on forewarnings, consider the various motives that warnings can induce, and evaluate how warnings affect thought about the issue in the appeal. In the reviewed research, some warnings informed recipients of a communicator’s intent to persuade by noting that “you will hear a tape recorded message... designed to change your mind” (Hass & Grady, 1975, p. 462). Others specified
the topic and position of the coming appeal, such as, "you will hear a tape recorded message advocating an increase in the New York City subway fare to 50 cents" (Hass & Grady, 1975, p. 462). Warning effects were then examined on people’s attitudes before they received the appeal as well as on people’s reactions to the appeal when it was delivered. At the end of the chapter, we consider the implications of forewarning research for the general phenomenon of resistance.

WARNINGS THAT INDUCE RESISTANCE: EVERYDAY INFLUENCE SETTINGS

In applied influence settings, forewarnings are used along with other interventions to instill resistance to subsequent counter-attitudinal attacks. Public health campaigns, such as those designed to prevent adolescent smoking, provide some of the clearest evidence of the effectiveness of these interventions. In the case of antismoking campaigns, the interventions typically are delivered in school and community settings to alert adolescents to social influences to smoke and to prepare them to resist these pressures. These programs might, for example, warn students of pressures to smoke from peers and from cigarette advertisers (Botvin & Kantor, 2000). Antismoking programs are multifaceted and typically include a variety of components in addition to forewarning. The programs might also provide information about smoking health risks and increase students’ self-efficacy. When structured appropriately, the programs seem to be effective in reducing adolescent smoking (e.g., Botvin & Kantor, 2000; Bruvold, 1993; Rooney & Murray, 1996). Furthermore, warnings appear to contribute to this impact. The most effective treatments do not simply provide factual information on cigarette use but instead use a variety of social influence methods, including in particular warning of impending influence and teaching of resistance skills (Bruvold, 1993; Rooney & Murray, 1996).

Political news coverage is another everyday context in which forewarnings have been used to induce resistance. Specifically, “adwatch” programs conducted by television news media provide critiques of political advertisements to help voters identify and resist misleading information (Jamieson, 1992; Jamieson & Cappella, 1997; McKinnon & Kaid, 1999). Adwatchers, like antismoking interventions, consist of warnings in conjunction with other components, such as the presentation of evidence supporting or refuting the arguments made in the ad. Warnings in adwatch coverage may take the form of spoken disclaimers alerting viewers that the information presented represents partisan advertisements and not actual news stories. Visual cues such as text overlays that label the ad’s content as “misleading” or “false” also warn recipients to be wary of the arguments presented in campaign ads. In some empirical tests, adwatch increased viewer skepticism and reduced ad effectiveness (Cappella & Jamieson, 1994; Leshner, 2001; Plaut & Louden, 1994). However, in other tests, adwatch yielded a “boomerang” effect in the form of increased recall of an ad’s arguments and increased favorability toward the candidate supported in the ad (Ansolabehere & Iyengar, 1996; Jamieson, 1992; McKinnon & Kaid, 1999; Plaut & Louden, 1994). These varying effects of adwatch are likely due to the varying content of the adwatch programs themselves. For example, the boomerang effect described above might occur when adwatch suggests that most of the claims made in a political ad are true or when they repeat the ad’s arguments with minimal refutation or other commentary (Jamieson & Cappella, 1997).

In summary, warnings have the potential in real-world settings to prepare people to resist temptations for unhealthy behavior and to recognize inaccuracies in political campaigns. Yet, because warnings have typically been treated as one component of applied intervention programs, it is unclear whether warnings alone have an impact or whether their intervention effects emerge in combination with other program features. Experimental research that has examined the independent effects of forewarnings is more directly informative.

YOU’VE BEEN WARNED . . .

Past narrative reviews of the experimental literature have concluded that forewarnings can have a variety of effects in addition to resistance. McGuire (1969) characterized the effects of forewarnings of messages that are intended to persuade as “frustratingly elusive.” He concluded that “there does seem to be a relationship begging to be found, and yet it seems to be hiding out in only certain cells of our experimental design” (p. 35). Other reviews have differentiated between types of warnings and concluded that forewarnings that convey just the persuasive intent of an appeal typically generate resistance, whereas forewarnings that specify the message topic and stance can generate resistance or susceptibility (Cialdini & Petty, 1981; Petty & Cacioppo, 1986). Eagly and Chaiken (1993) simply noted that these two types of warnings have widely varying effects depending on specific features of the warnings. More relevant to the focus of the present chapter is Petty and Wegener’s (1998) discussion of how warnings confer resistance by directing people’s thinking. They suggest that warnings of intent as well as warnings of topic and stance can reduce influence by negatively biasing people’s thoughts about the attitude issue and the subsequent appeal. In sum, although narrative reviews of the experimental literature have recognized that warnings sometimes confer resistance, they have also noted that this resistance is not an inevitable outcome of exposure to warning.

Given the range of warning effects noted in past reviews, it is no surprise that researchers have used a number of theoretical perspectives to explain how warnings function (see Cialdini & Petty, 1981). Most theoretical accounts consider how warnings instigate particular motivational states that direct recipients’ responses to an influence attempt. In this chapter we focus on three overarching motives that warnings might establish: (a) a conservative orientation to defend
existing attitudes and important self-identities, (b) a desire to maintain positive relations with others and to convey certain impressions, and (c) a wish to understand reality and to hold valid judgments (Chaiken, Giner-Sorolla, & Chen, 1996; Wood, 2000). In brief, we argue that whether a warning generates resistance depends on the specific motives that it elicits in a given influence context.

**Defense Motives.** Warnings of a counter-attitudinal attack can generate a defensive response by threatening recipients’ existing attitudes and their important identities, such as their personal integrity (Petty & Wegener, 1998; Zanna, 1993; McGuire & Millman, 1965). Warnings in public health campaigns are often designed for this purpose, with the intent of motivating people to self-protect through rejection of risky, health-threatening behaviors. In general, defensively motivated individuals evaluate the available information, including their existing attitudes, the warning, the source, the context, and the message if available, to express a judgment that best meets their defense goals of protecting their attitudes and identities (Chaiken et al., 1996).

Warnings that elicit defensiveness can have a variety of effects on thoughts about the attitude issue and the impending appeal. In response to warnings that threaten existing attitudes with an impending attack, people may selectively attend to attitude-consistent information, selectively encode or retrieve such information, counterargue the impending message, and bolster existing attitudes and beliefs (e.g., Cialdini & Petty, 1981; Eagly, Chboy, & Chaiken, & Shaw-Barnes, 1991; Petty & Cacioppo, 1977). Another form of defensiveness may emerge when warnings elicit reactance because they appear to limit people’s freedom to hold their favored position (Brehm, 1966; Hass & Grady, 1975). Then people may retain their original attitudes and resist the appeal in order to reestablish the threatened freedom, and they may do so with little thought about the message topic itself.

Some of the best evidence that warnings yield the resistance suggested by these defense theories emerged in Benoit’s (1998) meta-analytic synthesis of prior experiments on the effects of warnings on persuasion. This review of 12 studies revealed that participants who were warned prior to the receipt of a message were less persuaded than were participants who received a message without warning (mean r = .18). Benoit also reported that warning-induced resistance was uniform across several potential moderating variables. However, this uniformity could be due to the small number of studies in his sample, which would have provided limited opportunity to explore potential moderators.

Although resistance may seem the most likely outcome of warning-induced defensiveness, McGuire and Millman (1965) suggested that warnings that threaten personal identity and self-esteem can yield the—somewhat surprising—outcome of preemptive attitude change toward the message. These researchers reasoned that warnings can make people anxious about being gullible or easily influenced when people believe that they are ultimately likely to be persuaded by the message. Then, the best strategy to preserve self-esteem is to shift preemptively toward the message to minimize its apparent impact. McGuire and Millman (1965) recognized further that preemptive shifts can both defend the self-concept and, when made public, secure the favorable impressions of others by not appearing to be influenced. However, they emphasized self-related motives such as maximizing self-esteem and minimizing discrepancies between real and ideal selves. In this spirit, we interpret such attitude change as a defensive maneuver intended to protect the self against being gullible rather than a self-presentational maneuver.

**Impression Motives.** Under certain conditions, warnings can make people concerned about the impressions they convey to others. Forewarnings elicit impression motives when they alert recipients to the interpersonal consequences of expressing an opinion (Cialdini, Levy, Herman, & Evenbeck, 1973; Cialdini, Levy, Herman, Kozlowski, & Petty, 1976; Lundgren & Pruslin, 1998). For example, antismoking interventions that warn teenagers of impending peer pressure to smoke can allow teens to develop response strategies that minimize social rejection. In general, impression-motivated individuals evaluate the available information to express a judgment that best conveys the desired impression (Chaiken et al., 1996). Research investigating impression motives has typically warned people of an impending discussion with another who holds opposing views (e.g., Cialdini et al., 1973; Cialdini et al., 1976). People then may express moderate attitudes because a moderate stance makes them appear broad-minded and responsive and allows them to use arguments from both sides of an issue to support their position in the impending discussion.

**Validity Motives.** Warnings can make people concerned about the validity of their judgments when the warning indicates that the impending message selectively favors a certain position or is biased in some way. For example, warnings in the “adwatch” programs were designed to heighten viewers’ awareness of biased information in political campaign ads. People motivated by accuracy concerns evaluate the available information in search of a judgment that best represents objective reality (Chaiken et al., 1996). Although experimental research has not evaluated this potential effect of warnings in much detail, research on bias correction has addressed judgment validity (Wegener & Petty, 1997). That is, when informed that a certain stimulus may have a biasing effect that threatens judgment validity (e.g., an especially attractive source), people have been found to correct for this effect and shift their attitudes to counteract the presumed bias.

In sum, past experimental research has documented a variety of effects of warnings, with some studies indicating that warnings induce resistance to an impending appeal and others indicating that warnings increase acceptance of the appeal. To explain these varying effects, theories of warning impact have identified a variety of motives likely to be instigated by warnings. Specifically, warnings that threaten people’s existing attitudes are likely to instigate a defen-
sive orientation and resistance to protect those attitudes. Warnings that threaten people's self-concepts by suggesting that they are vulnerable to influence are likely to instigate a defensive orientation and preemptive attitude change in order to reduce the apparent impact of the appeal. Warnings that focus people on the impressions they convey through their response to the message are likely to yield moderation toward a neutral, easily defensible position. To evaluate the plausibility of each of these interpretations, we conducted a meta-analytic synthesis of forewarning impact (Wood & Quinn, 2003).

Wood and Quinn's (2003) Research Synthesis

We conducted a comprehensive assessment of the direction and magnitude of forewarning effects on attitudes and examined how warning impact varies with the experimental conditions represented in forewarning studies. To accomplish this, we needed a broader sample of studies than Benoit's (1998) earlier investigation, which was limited to studies that evaluated warning impact on reactions to a persuasive appeal. Thus, we included experiments that assessed recipients' attitudes following the warning but prior to receipt of the appeal. These studies reveal how simply expecting an appeal influences attitudes. Overall, then, our review provided an estimate of warning impact prior to a message as well as warning impact on message persuasiveness.

We located eligible studies through a variety of mechanisms, including computerized databases (e.g., PsyInfo, ERIC, Dissertation Abstracts) and reference lists from previous reviews (e.g., Benoit, 1998; Cialdini & Petty, 1981). A total of 37 research reports were located that examined the effects on attitudes of warnings of persuasive appeals. Of these, 19 assessed attitudes after the warning but before message delivery. Several of these contained multiple experiments, yielding 25 separate studies in the pre-message sample. Eighteen reports assessed attitudes following delivery of both the warning and the message. Several of these also contained multiple experiments, yielding 23 separate studies in the post-message sample.

We analyzed pre-message and post-message samples separately to capture the unique effects of warnings in each paradigm. Prior reviews have given little reason to suppose that warning impact varies with the delivery of the message. For example, Cialdini and Petty (1981) claimed that the direction of forewarning impact "cannot be explained through an examination of whether . . . effects have been assessed in terms of pre- or postcommunication attitude change" (p. 221). Thus, initially we had few expectations that timing of assessment would influence warning impact.

To calculate effect sizes, we compared the attitudes participants expressed after the warning with the attitudes of a nonwarned comparison group. In the pre-message sample, the attitudes of both the warned and the nonwarned groups were assessed prior to the message. In the post-message sample, the attitudes of both the warned and nonwarned groups were assessed after the message had been delivered. Effect sizes were calculated for each study in terms of $d$, the mean difference between the warned and nonwarned groups' attitudes, divided by (for between-groups comparisons) the pooled standard deviation or (for within-group comparisons) the standard deviation of the difference between paired observations (Hedges & Olkin, 1985). Because the position of the impending appeal was always on the opposite side of the neutral point from participants' initial positions, agreement shifts toward the appeal were in the same direction as moderation shifts toward the scale midpoint. Thus, effect sizes were given a positive sign to signify change toward the appeal and/or the scale midpoint and a negative sign to indicate change away from the appeal and/or polarization.

### Table 10.1

<table>
<thead>
<tr>
<th>Study groupings</th>
<th>$k$</th>
<th>Mean weighted effect size ($d$)</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attitudes after warning but before message</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full sample</td>
<td>20</td>
<td>0.37</td>
<td>0.29/0.46</td>
</tr>
<tr>
<td>Message canceled following warning</td>
<td>7</td>
<td>0.08</td>
<td>-0.19/0.25</td>
</tr>
<tr>
<td>Topic-only warnings</td>
<td>3</td>
<td>0.11</td>
<td>-0.25/0.46</td>
</tr>
<tr>
<td>Thought listing followed warning</td>
<td>4</td>
<td>-0.02</td>
<td>-0.19/0.15</td>
</tr>
<tr>
<td>Distraction following warning</td>
<td>2</td>
<td>0.61</td>
<td>0.27/0.95</td>
</tr>
<tr>
<td>Studies that manipulated outcome-relevant involvement of message topic Low</td>
<td>3</td>
<td>-0.32</td>
<td>-0.61/-0.03</td>
</tr>
<tr>
<td>Attitudes following warning and message</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full sample (compared with nonwarned controls who received message)</td>
<td>17</td>
<td>-0.38</td>
<td>-0.48/-0.28</td>
</tr>
<tr>
<td>Warning impact compared with no-message control conditions</td>
<td>7</td>
<td>0.22</td>
<td>0.06/0.37</td>
</tr>
<tr>
<td>Distraction between warning and message</td>
<td>3</td>
<td>0.20</td>
<td>-0.10/0.50</td>
</tr>
<tr>
<td>Distraction plus time delay between warning and message</td>
<td>3</td>
<td>0.05</td>
<td>-0.13/0.23</td>
</tr>
<tr>
<td>Thought listing between warning and message</td>
<td>3</td>
<td>-0.64</td>
<td>-0.94/-0.34</td>
</tr>
<tr>
<td>Studies that manipulated outcome-relevant involvement of message topic Low</td>
<td>6</td>
<td>-0.96</td>
<td>-1.25/-0.66</td>
</tr>
</tbody>
</table>

$k$ = the number of study findings, and thus the number of effect sizes, in each mean effect size calculation.

The $d$s are standardized difference scores representing shifts in the attitudes of forewarned participants in comparison with attitudes of those who were not warned. All mean $d$s were computed with each effect size weighted by the reciprocal of its variance; a procedure that gives more weight to effect sizes that are more reliably estimated (Hedges & Olkin, 1985). Negative effect sizes indicate that warnings were associated with greater resistance to the position advocated in the pre-message compared with controls. Positive effect sizes indicate that warnings were associated with increased acceptance of the position advocated in the message in comparison with controls.
Forewarnings That Induce Resistance. In general, warnings elicited resistance most consistently when attitudes were assessed following the appeal (see bottom panel of Table 10.1). That is, people who received a warning before a message were less persuaded by the message than those who received a message without warning. In addition, this warning-induced resistance did not appear to emerge from a reactance-like desire to reestablish freedom to hold existing attitudes (see Brehm, 1966). In both the pre-appeal and the post-appeal synthesizes, warnings that specified an intent to persuade recipients, and thus that specifically attacked their freedom, did not elicit greater resistance than warnings that did not mention such an intent. The synthesis thus provided little evidence that warning impact emerged from reactance-like motives.

Evidence that forewarning-induced resistance was a thoughtful process emerged from studies that instituted a challenging cognitive task between the warning and the appeal (e.g., solving mathematical and verbal problems). Such tasks are distracting and reduce people’s ability to build cognitive defenses against the impending appeal (Eagly & Chaiken, 1993). Despite the general finding that warning-induced resistance emerged when recipients’ attitudes were assessed after the warning and receipt of the message, no effect of warning emerged in post-appeal studies that distracted participants between the warning and the appeal (see Table 10.1). Distracted individuals apparently were unable to muster the cognitive support to counterargue the impending appeal and were no more likely than their nonwarned counterparts to resist the message.

Additional evidence of the thoughtful nature of warning-induced resistance emerged with the studies that presented warnings on topics that were highly involving in the sense that they concerned immediately important outcomes. For these kinds of issues, recipients generally are highly motivated to scrutinize message-relevant information (Johnson & Eagly, 1989). Warnings apparently biased this processing in a negative direction. Thus, warnings on involving topics yielded strong evidence of resistance in both the pre-appeal and the post-appeal paradigms (see Table 10.1). Further suggesting the role of thought in generating attitudinal resistance, studies that explicitly instructed recipients to report their thoughts following the warning obtained minimal attitude change toward the impending appeal (see Table 10.1). Thus, warnings were especially likely to yield resistance when they were given in conjunction with involving messages or direct instructions to list thoughts. Under these conditions, the threat posed by the warnings apparently biased recipients’ processing and resulted in attitudinal resistance.

Studies that instituted a time delay after the warning and thereby gave participants the opportunity to deliberate offer additional insight into the thoughtful nature of warning-induced resistance. Forewarning researchers often incorporated time delays into their experimental procedures because delays supposedly allow participants the opportunity to construct counterarguments prior to message receipt (Cialdini & Petty, 1981; Jacks & Devine, 2000). However, the results of studies that used a time delay did not differ from those that did not, indicating that in general time delays were not associated with resistance (see similar results in Benoit, 1998). To investigate further the effects of time delay, we separately evaluated a subset of three studies that are commonly cited in narrative reviews as evidence that delays increase resistance (e.g., Cialdini & Petty, 1981). These three studies directly manipulated delay and thus provide direct evidence of the effectiveness of this technique (Freedman & Sears, 1965; Hess & Grady, 1975; Petty & Cacioppo, 1977, Experiment 1). Interestingly, all three studies used messages on highly involving topics, and therefore recipients should already have been motivated to think carefully about the message topic. In this subset of studies, a delay following the warning yielded greater resistance than when warnings were presented with no delay, suggesting that the time delay provided the opportunity to engage in thought. Thus, it seems likely that the null effects for time delay across the whole sample of studies were due to the primarily uninvolving topics typical of this sample. In general, the pattern of findings suggests that although delay provides the opportunity to think, this opportunity has little impact unless participants are already motivated to undertake such cognitive activity—by, for example, an involving topic.

The thoughtful resistance induced by warnings is similar to that found with research on inoculation, in which people who receive weak versions of an attacking message have been found to be protected against future threatening messages. Inoculation yields resistance because it motivates and enables cognitive refutation of counter-attitudinal appeals (Pflau, 1996; McGuire, 1964). The thoughtful resistance generated by forewarnings also bears some similarity to the cognitive processing mechanisms identified in research on motivated reasoning (Kunda, 1990). People who are motivated to arrive at a judgment that affords them the most favorable outcomes appear to engage in a biased memory search and to construct congenial beliefs in order to support that judgment. This cognitive bolstering allows people to resist evidence that counters their preferred positions (Kunda, 1990).

In general, the cognitive elaboration underlying warning-induced resistance should render recipients’ attitudes relatively persistent and predictive of future behavior (Petty, Haugtvedt, & Smith, 1995). Thus, post-message warnings likely generated strong attitudes that would remain relatively resistant to future attack.

Forewarnings That Shift Attitudes Toward the Message. We suggested earlier in this chapter that defensive responses to warnings do not always yield resistance. As McGuire and Millman (1965) argued, warnings that threaten recipients’ identities can induce movement toward impending appeals. These preemptive shifts can minimize the threat of being gullible and changing one’s attitude. Supporting this idea, preemptive change was the mean finding across the whole sample of studies when attitudes were assessed prior to message delivery (see top panel in Table 10.1). It is interesting to note that the shift toward the appeal was comparable in size to the resistance effect that emerged across the post-appeal sample of studies, suggesting that the magnitude of warning impact was uniform across the two paradigms, although opposite in direction. Furthermore, as would be expected if participants were responding to the
threat of eventual attitude change, the preemptive attitude change was especially marked when the communication was reputed to be highly persuasive or was delivered by an expert source—conditions that should increase the likelihood that people believe they will be persuaded.

The pattern of pre-appeal attitude change appears to have been a relatively superficial response accompanied by minimal issue-relevant thought. Specifically, in the two pre-appeal studies that distracted participants after the warning was presented but before their attitudes were assessed, participants’ attitude change was similar to that obtained when there was no distraction (see Table 10.1). Even though distraction presumably inhibited thought, it did not reduce the amount of movement toward the message. Thus, change toward the appeal plausibly reflected heuristic processing and recipients’ use of rules such as “persuasive messages, expert sources . . . must be correct.” Participants likely selected the rule that best met their defensive goals of maintaining a favorable self-view. Presumably, participants were unaware that using this rule generated attitude change; any awareness of preemptive agreement should reduce its effectiveness in defending self-esteem.

Additional evidence that warning-induced shifts toward the appeal were based on minimal issue-relevant thought emerged in the few studies that alleviated the threat of the impending appeal by canceling the message (see Table 10.1). When the message was canceled, preemptive attitude change disappeared, and participants’ attitudes returned to their original positions. Thus, preemptive change represented an “elastic shift” or temporary response that emerged only while participants were anticipating the appeal (Cialdini et al., 1973). The elastic nature of attitude change provides important support for our argument that shifts toward the message emerged in response to defensive motives. Attitudes based solely on informational cues (e.g., expert sources, purportedly persuasive appeals) should have withstood cancellation because these cues would have remained available after the message threat was removed—and these cues might even have been strengthened by mention of the message in the cancellation manipulation. Additional evidence that informational processes alone did not account for anticipatory shifts toward the appeal was provided by studies that included an additional control condition in which participants were informed of the impending appeal but did not expect to be exposed to it themselves. Simply being informed about the existence of a counter-attitudinal appeal was not sufficient to change attitudes. Thus, the overall pattern of findings is consistent with the idea that anticipatory attitude change was a strategic response to the warning-induced threat to the self-concept.

One of the most interesting findings of the review is that warnings yielded very different patterns of attitude change before and after message delivery. Prior to the appeal, the threat represented in the warning generated change toward the message position, but upon receipt of the appeal it generated resistance (as we noted in our discussion of the post-appeal findings; see top panel of Table 10.1). These divergent effects of warning can be attributed to differences in the apparent target of the warning-induced threat prior to and during the appeal. Prior to the appeal, people apparently were motivated to defend a view of themselves as not being easily influenced. During and after receipt of the appeal, they apparently were motivated to defend their attitudes against the attacking message. This apparent shift in focus from defending the self to defending one’s attitudes that accompanied delivery of the appeal could reflect a variety of factors. One is that people may not have been able thoughtfully to resist an appeal before learning what it said. In general, the idea that people can reduce threat in influence settings through both agreement and rejection has parallels in research on fear appeals. Fear-arousing components of messages have been found to increase persuasion or increase rejection, depending on such factors as the intensity of the threat and the available coping strategies (Gleicher & Petty, 1992; Baron, Logan, Lilly, Inman, & Brennan, 1994).

**Conscious Awareness of Warning Impact.** The research we reviewed on forewarnings did not address whether warnings operate consciously or nonconsciously. It seems plausible that some amount of conscious thought was involved in warning-induced resistance, given that resistance emerged from bolstering existing attitudes and refuting the advocated position. However, this thoughtful analysis might be undertaken as a deliberate defensive response or it might be spurred directly by perceptions of a communicator’s persuasive intent with little explicit recognition of the need for defense (Dijksterhuis & Bargh, 2001). Thus, people are not necessarily aware that they are motivated to defend their attitudes and resist the appeal in response to the warning. In contrast, we can be reasonably sure that preemptive shifts toward the appeal were enacted without conscious awareness. Our explanation that anticipatory shifts were undertaken to preserve self-esteem implies that the strategy operates outside of consciousness—any awareness of one’s vulnerability to a persuasive attack and preemptive change to bolster the self would likely undermine the effectiveness of this strategy. In general, then, warnings may sometimes yield resistance when people consciously decide to defend their views or do so nonconsciously. However, preemptive change toward the message most likely occurs with limited conscious awareness.

**Impression Motives Instigated by Forewarnings**

The suggestive evidence from our review that warnings elicit a defensive response can be contrasted with the possibility that warnings sensitize people to the impressions they convey to others (Cialdini et al., 1973; Cialdini et al., 1976). As we mentioned earlier in this chapter, warnings can instigate impression motives when they orient people to appear open-minded to others and to hold attitudes that will be defensible in discussions. To convey these favorable impressions, warned recipients may shift to neutral, mid-scale positions that appear unbiased and that allow them to use a range of arguments to explain their position.
Because the motivational states instigated by forewarnings typically were not assessed directly in the research, it is difficult to ascertain the specific defense- or impression-motivated state that warnings might have established in a given experimental context. As reviewers, we are functionally in the same position as the study authors, forced to infer the participants' motivational states from the pattern of attitude change findings. This failure to directly measure motives extends beyond the warning literature. Previous meta-analytic syntheses of persuasion and social influence also noted that the reviewed studies did not obtain direct measures of the motivational states instigating attitude change (e.g., Johnson & Eagly, 1989; Wood, Lundgren, Ouellette, Busceme, & Blackstone, 1994). Thus, in reviewing the warning literature we were unable to evaluate directly the possibility that attitudes shifted in response to specific motivational concerns instigated by a warning. We instead evaluated whether the pattern of attitude change outcomes was consistent with the idea that warnings elicited specific motives.

To evaluate impression motives, we compared the attitudes participants expressed when they expected to report their post-appeal attitudes publicly to others versus privately on a questionnaire. If attitude change is a relatively superficial attempt to convey certain impressions to others, then it should emerge primarily in public settings in which people can garner the social benefits of conveying particular impressions (Cialdini et al., 1973; Cialdini & Petty, 1981). However, across all studies in the review, attitude change was not greater with expectation of public rather than private attitude judgments. Additional evidence against an impression-motivated interpretation of attitude change prior to the appeal was provided by studies with warnings that specified only the topic and not the position in the impending appeal (see Table 10.1). Because impression-motivated attitudinal moderation requires just knowledge of the topic of the appeal, moderation shifts would be expected to emerge with topic-only warnings. However, we observed significant attitude shifts only in studies in which warnings indicated that the appeal would be counter-attitudinal, either by stating the message position or an intent to persuade. Thus, it appears that recipients agreed with the specific position in the impending appeal and were not engaging in impression-motivated moderation of their attitudes in expectation of any position that they might encounter on a topic.

Finally, we note that parsimony is an additional reason to favor defense over impression accounts of warning impact. Warning-induced defensiveness can explain both the acquiescence effects apparent prior to the appeal and the resistance effects that emerged following the appeal and when the message topic was highly involving. Thus, forewarnings of an impending counter-attitudinal attack appear to have motivated recipients to defend their existing attitudes and identities, rather than to appear in a favorable manner to others.

Although there was little overall evidence that impression concerns were responsible for the attitude shifts prior to the appeal, it remains plausible that certain types of warnings given in certain types of contexts can heighten concerns about the impressions conveyed to others. Specifically, Cialdini et al. (1973) proposed that warnings would be especially likely to heighten impression concerns when they inform recipients of (a) an impending discussion on (b) an uninviting topic with (c) a person who holds an opposing opinion. Two studies in Wood and Quinn's synthesis established these conditions in certain cells of their experimental designs (i.e., Cialdini et al., 1973; Deaux, 1968). Consistent with an impression interpretation of attitude change, these specified conditions participants warned of a discussion on the message topic with another person shifted their attitudes toward midrange and the impending partner's position. These shifts were greater than those observed when warned participants expected to indicate their attitudes privately on a questionnaire. Thus, warnings of an impending discussion of interaction may operate differently than the warnings of an attack message in that the former sensitizes people to the impressions they convey to others whereas the latter highlights defensive concerns.

Although these two discussion warning studies in our synthesis found that people moderated their attitudes to an impending discussion, it is also possible that impression motives can lead to resistance under certain circumstances (Fitzpatrick & Eagly, 1981). Resistance is likely when discussions are framed as debates or competitions and a favorable impression can be attained by prevailing over a discussion partner. People also may resist when their initial attitudes are known by the discussion partner and changing their views would give the impression of inconsistency or compliance. In general, then, as with warnings that instigate defensive concerns, warnings that elicit impression motives are plausibly associated with a variety of attitudinal outcomes.

RESISTANCE AS AN OUTCOME OF FOREWARNING

The research reviewed in this chapter illustrates several attitudinal effects that result from warning of an impending counter-attitudinal appeal. Consistent with the expectations of practitioners who use warnings to forestall recipients and help them withstand a persuasive attack, forewarnings often produced defensive resistance to the message position. However, under certain conditions the same forewarning treatment had the opposite effect, eroding rather than enhancing resistance to a counter-attitudinal appeal. In this section of the chapter, we draw on the forewarning literature to develop answers to the general questions of what resistance is and when it is likely to occur. Our position is foreshadowed by the title of this section.

Formal definitions of resistance are difficult to find in the attitudes literature (see Jacks & O'Brien, this volume). Whereas some authors assert that resistance is not merely the inverse of persuasion (McGuire, 1964; Knowles, Butler, & Linn, 2001), few have defined exactly what resistance is (with notable exceptions of this volume; Petty & Wegener, 1998; Brehm, 1966). Resistance potentially can be treated as the motive to reject influence attempts and retain one's own views, as the process through which people reject and retain, and as
a potential outcome of influence. As we explain, the pattern of attitude change that emerged in studies of forewarning suggests that resistance does not correspond to a particular motive or a particular process of rejection; it is best understood as an outcome of an influence attempt. In addition, we note that because the forewarning literature does not address another possible conceptualization of resistance—as a quality of an attitude or a person—we will not discuss this conceptualization here (see Brifol, Rucker, Tormala, & Petty, this volume).

Is Resistance a Motive? Reactance theory provides a classic perspective on resistance to influence as an underlying motive (Brehn, 1966). According to this analysis, people desire to maintain freedom and, in influence contexts, to assert their own positions and oppose the appeal. Similarly, Jacks and O'Brien (this volume) define resistance as a motive to withstand an influence attempt. Yet, the warning literature illustrates the difficulties that emerge when resistance is defined as a motive. First, it is evident that this motive does not correspond to reactance. In the research synthesis, resistance was not heightened in circumstances that should have increased participants’ reactance and desire to assert freedom through message rejection (i.e., when the warning mentioned an intent to persuade). Although a reactance-like motive may lead to resistance in some circumstances, it did not appear to account for resistance generated by forewarnings.

Instead of a general motive to resist, the results from forewarning studies suggest that multiple specific motives directed resistance and agreement (see also Wegener, Petty, Smoak, & Fabrigar, this volume). Specifically, resistance was found when warnings appeared to instigate defense motives in recipients. That is, resistance was pronounced when warnings apparently threatened participants’ attitudes, as occurred when the warning topic was highly involving. That this specific defense motive did not represent a general desire to resist was evident in the pre-appeal literature, in which warning-induced threats to identity generated defensiveness in the form of preemptive opposition. Other research has indicated that, in addition to defense motives, warning-induced resistance can stem from a motivational concern with conveying specific impressions to others (e.g., Fitzpatrick & Eagly, 1981). That is, people resist when they wish to convey the impression of prevailing over an opponent in a debate or when their initial attitudes are known and change would signal compliance. Also, accuracy goals potentially could lead to resistance under some circumstances. We speculate that people motivated to be accurate will not resist a highly valid, credible message but will strongly reject a biased or one-sided position.

In summary, resistance in the forewarning literature did not emerge as a global reactance motive. Instead, resistance seemed to reflect situation-specific defense and impression motives, and potentially could reflect accuracy motives.

Is Resistance a Process? Resistance as a process represents a negatively biased conscious or nonconscious reaction against an influence attempt that leads to rejection of the appeal. The classic example of resistance as a process involves thought generation as people counterargue appeals and bolster their existing views. Given that these mechanisms were critical to warning-induced resistance, it is tempting to define resistance in terms of this process. As we noted, when participants in post-message studies were distracted and apparently could not generate thoughts, warnings had little impact. Conversely, when participants were instructed to list their thoughts following the warning, they were especially successful at resisting the appeal.

Despite the evidence from forewarning that resistance emerged from a thoughtful evaluation of the impending appeal, other research has demonstrated that resistance implicates a variety of processing mechanisms. For example, resistance can involve emotional responses, as when feelings of irritation and other negative affective cues signal disagreement with the message (e.g., Zuckerman & Devine, 1996). The process of resistance also can involve reliance on heuristic rules that suggest rejection of a message (Zuckerman & Chaiken, 1998). All of these processes involved in resistance can result in rejection of an appeal. We speculate that less thoughtful rejection of appeals will be found with warnings that generate a stronger defensive reaction than is typical of the warnings in the literature we reviewed. We base this idea on studies of strong fear-inducing appeals that appear to elicit minimally thoughtful resistance processes, including defensive avoidance and discounting of the source and message (Jepson & Chaiken, 1990; Gleicher & Petty, 1992).

In summary, resistance mechanisms include a variety of thoughtful and less thoughtful processes as well as emotional reactions. At a general level, these mechanisms are not just involved in resistance and rejection of appeals but also are components of persuasion processes, as people thoughtfully or less thoughtfully accept messages and respond to them in emotionally positive ways.

Resistance Is Best Understood as an Outcome. Attitudinal outcomes yield a clear, unambiguous indicator of resistance. Attitude outcomes are the gold standard for resistance: In our view, resistance does not occur without attitudinal evidence. In the warning literature, resistance is evident when the attitudes reported by warned participants are less favorable than those held by nonwarned participants. Theories explaining warning-induced resistance identity how this outcome is motivated by defense, impression, or accuracy concerns and how it is generated through various modes of processing (e.g., thoughtful or less thoughtful).

Our review of forewarning research identified several patterns of resistance when warned participants were compared with other, nonwarned ones. Prior to message delivery, when warnings concerned highly involving issues, resistance emerged as a “boomerang effect” or attitude shift away from an impending appeal. Boomerang shifts also emerged following message receipt when warned individuals had listed their thoughts prior to the appeal (and were compared with nonwarned controls who did not receive the message). Apparently, these experimental contexts spurred such persuasive cognitive bolstering of existing
views and counterarguing of the appeal that people moved away from the expected position.

Resistance as attenuated agreement emerged following message delivery when warned participants shifted toward the message, but shifted to a lesser extent than nonwarned participants. Because messages in general were persuasive, the resistance-inducing effects of warnings reduced but did not eradicate message impact. In these experimental contexts, warnings apparently negatively biased message evaluation and thus decreased the amount of attitude change. It is interesting to note that although we are defining these outcomes as "resistance," in actuality the mean attitudinal effect even after a warning was a shift toward the message position.

Resistance as maintenance of original views given pressure to change was apparent prior to the appeal. That is, participants who listed their thoughts following the warning apparently engaged in bolstering of their own views and counterrarguing of the impending appeal. Thus, they retained their views and did not demonstrate the shifts toward the appeal evidenced by participants who did not list their thoughts. In general, then, resistance outcomes generated by warnings were reflected in a variety of patterns of change, including boomerang shifts away from an impending appeal, reduced amounts of change toward the appeal, and no attitude change.

**Outcomes Not Representative of Resistance.** Failure to change one's attitudes does not always signal resistance. Resistance occurs when individuals face pressure to change their attitudes. The three examples of resistance from the warning literature—a boomerang shift, reduced message persuasiveness, and attitudinal stability—all emerged in response to influence pressures that were in fact successful at changing the attitudes of nonwarned participants and other comparison groups. In contrast, attitude stability due to lack of pressure can be found when people are not exposed to any persuasive information, when they are exposed to attitude-congruent appeals, and when they are exposed to counter-attitudinal appeals with very weak arguments. Messages with weak arguments often link the attitude object to unfavorable or to neutral outcomes, and consequently information of this kind does not present much pressure to adopt the advocated position (for an opposing view of resistance, see Johnson, Smith-McLallen, Killeya, & Levin, this volume).

**APPLYING FOREWARNING RESEARCH**

The research reviewed in this chapter will be helpful to practitioners using forewarnings to instill resistance to influence. Practitioners' efforts will most likely be successful when they understand the motives instigated by their warning interventions and ensure that the context in which the warning is delivered allows recipients to satisfy those motives through resistance rather than acquiescence. To generate warning-induced resistance of the form obtained in the reviewed studies, practitioners should use warnings that (a) threaten attitudes, and (b) foster issue-relevant thought by highlighting the involving aspects of the issue or by explicitly instructing people to consider the issue and the impact of the message's counter-attitudinal position. Warnings should also be delivered in a context that is relatively free of distractions and thereby allows for systematic processing of relevant information.

Our findings also suggest that forewarning can be a persuasion tactic. Practitioners may attempt to use warnings to threaten recipients' identities rather than their attitudes in order to generate preemptive agreement (see also Sherman, Crawford, & McConnell, this volume, for a discussion of how future regret may generate compliance). However, in the reviewed research this acquiescence was a temporary and context-dependent response that disappeared when the threat was removed and the message was canceled. Thus, the compliance effects of warnings may be only short lived and may not represent enduring attitude change.

Warnings that threaten people's identities may not always lead to acquiescence, and may instead yield resistance when the identity involves an important social group. That is, when people identify strongly with a group, they are likely to be motivated to hold attitudes consistent with those of typical group members and inconsistent with those held by the outgroup. Following this logic, warnings are likely to instigate resistance when they associate an attitude position with an outgroup or derogated minority (e.g., "...keep in mind, this is what the outgroup would have you believe"). Similarly, resistance is likely when warnings claim attitude positions are typical of a valued ingroup (e.g., "...we are the ingroup and we believe x, not y"). The effects of specifying social identities in warnings could emerge because the identities act as simple persuasion cues (Pratkanis & Turner, 1996) or because the identities inscringest thought about the message position (Pool, Wood, & Leck, 1998; Wood, Pool, Leck, & Purvis, 1996). The social identity effects of warnings are at present speculative and remain to be explored in subsequent research.

It would be premature for influence practitioners to attempt to instill resistance through warnings that sensitize people to the impressions they convey to others. The findings of our review provide only preliminary hints of the ways that warnings can enhance or reduce resistance by raising impression concerns. Attitudinal moderation toward midscale emerged in analyses on the two studies presenting warnings of an impending discussion with another person who held a counter-attitudinal position on a low involving issue. Furthermore, the conditions that lead to impression-motivated resistance are not well understood; this effect only emerged in one study that could not be included in the synthesis because it did not provide sufficient data to calculate effect sizes (Fitzpatrick & Eagly, 1981). Thus, an impression analysis of warning impact awaits further research.

Even less evidence is available about resistance generated by concerns with position accuracy. Our speculation that warnings could heighten accuracy concerns was based on research investigating techniques other than forewarnings.
Specifically, research on bias correction has demonstrated that a perceived threat to the objectivity of one’s judgment can lead to attitude shifts in a direction that resists the potential bias (Wegener & Petty, 1997). However, it is unclear whether warnings of potential bias have a general resistance effect. Much of the research on bias correction has been conducted with issues of low involvement (e.g., rating the desirability of vaccinating in various locales; rating the attractiveness of various celebrities; Wegener & Petty, 1997), and it may be that warnings of potential bias are motivating primarily when people do not have a strong initial attitude or an initial commitment to a particular position. Additional research will be needed to address the overall impact of forewarnings of potentially biased appeals along with the boundary conditions of such effects.

CONCLUSION

We have argued in this chapter that first warned may indeed be forearmed, but that forewarning does not always involve resistance. Warnings of impending influence can generate resistance when they orient recipients to consider the threat to their existing attitudes and undertake a cognitive defense. and when the warnings are presented in a context that does not distract people from careful thought. In contrast, forewarnings can increase susceptibility to persuasion when they focus recipients on the self-related implications of being influenced. That is, when recipients are concerned about being gullible and losing integrity, they may preemptively agree with appeals in order to minimize eventual change. Effective use of forewarnings in applied influence settings requires that practitioners consider the motives instigated by their interventions, the cognitive consequences of such events, and the contextual features that influence recipients' responses. Practitioners with an understanding of warning effects have at their disposal a means of arming individuals to resist harmful messages and thereby to foster desirable outcomes such as improved public health and more accurate political knowledge.

AUTHOR NOTE

Preparation of this chapter was supported by a grant from the National Institute of Health to the second author (1R01MH619000-01). Correspondence concerning the chapter should be directed to Jeffrey Quinn or Wendy Wood who are now at Duke University, Department of Psychology: Social and Health Sciences, Durham, NC 27708, Electronic mail j.quinn@duke.edu or wwood@duke.edu.

REFERENCES


