October 1975

An Emotion-Attribution Approach to Moral Behavior: Interfacing Cognitive and Avoidance Theories of Moral Development

Richard A. Dienstbier  
*University of Nebraska-Lincoln, rdienstbier2@unl.edu*

Donald Hillman  
*University of Nebraska - Lincoln*

John Lehnhoff  
*University of Nebraska - Lincoln*

Judith Hillman  
*University of Nebraska - Lincoln*

Maureen C. Valkenaar  
*University of Nebraska - Lincoln*

Follow this and additional works at: [http://digitalcommons.unl.edu/psychfacpub](http://digitalcommons.unl.edu/psychfacpub)  
Part of the *Psychiatry and Psychology Commons*

[http://digitalcommons.unl.edu/psychfacpub/254](http://digitalcommons.unl.edu/psychfacpub/254)

This Article is brought to you for free and open access by the Psychology, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Publications, Department of Psychology by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.
An Emotion-Attribution Approach to Moral Behavior: Interfacing Cognitive and Avoidance Theories of Moral Development

Richard A. Dienstbier, Donald Hillman, John Lehnhoff, Judith Hillman, and Maureen C. Valkenaar

University of Nebraska–Lincoln

Abstract: A theory is presented concerning the impact of attributions about the causes of emotional responses as they influence self-control in temptation situations. Research is reviewed indicating a high level of adult sensitivity to external influence in making such causal attributions. Two studies are presented in which the post transgression emotions of second-grade children are labeled shame (because of being found out) or guilt (due to the transgression itself); when a similar situation was subsequently represented as safe from detection, shame-condition children transgressed 60–80% more than guilt-condition subjects. It is suggested that emotional arousal elicited in temptation situations because of past punishment or options that are inconsistent with the self-image is necessary for inhibition but not sufficient unless attributed to a relevant cause. The literature on the relative effectiveness of moral socialization techniques is discussed with respect to the theory, and the relevance to cognitive dissonance and to overjustification approaches to motivation is discussed. An integration of social-learning and cognitive-developmental theories is approached through explicating the translation of moral decision into behavior by focusing on the ways that cognition may exert partial control over the impact of less finely differentiated emotional response, allowing cognitive overrides of contradictory emotional dispositions without eliminating the emotion.

Although both approaches are supported by the data of a great many studies, the two approaches remain essentially unintegrated. We seek to explicate some of the important steps in the translation of moral decision into moral behavior and to apply that analysis to an explanation of why certain moral socialization techniques are more conducive to internalization than others. Our theoretical perspective is presented in the early sections of the article; two supporting studies follow. The emphasis of the final section is on the relevance of the theory to traditional approaches to morality and to the socialization literature.

The theory can be stated in outline as follows: Behavior associated with self-control in the context of temptation is heavily influenced by negative emotional states usually characterized as anxiety, fear, guilt, and/or shame. This emotional mediation of moral behavior continues into adulthood, when...
the resultant behavior may relate to higher orders of moral judgment as defined by modern cognitive approaches to morality. Although the association of these negative emotional states with decisional and behavioral processes may take place in part through mechanisms of conditioning and learning specified by traditional avoidance learning theorists (e.g., Mowrer, 1950), these associations depend heavily on the causal attributions that are made about the source of the negative emotions during socialization experiences. Social and situational influences (often temporary) play an important role in forming causal attributions; different physical and verbal socialization techniques particularly provide different information relevant to the child’s causal attributions. Later, in the face of temptation, the impact of the emotion response on behavior will be heavily influenced by the still malleable beliefs held about the causal origins of the emotional response. When the individual believes detection of transgression is not possible, a negative emotional response in the face of temptation is necessary but not sufficient for the inhibition of transgression. In order for that emotional arousal to serve an inhibitory function, the individual must identify his emotional discomfort as due to a relevant cause, such as the transgression per se (and the implications of the transgression for self-image, etc.), rather than as due to an irrelevant cause, such as the fear of punishment. With maturity, control of emotional attributions passes from the socializing agent to the individual and becomes related to the level of moral development, although situational and social cues continue to play an important role.

Emotion Attribution

The work of Schachter and Singer (1962) demonstrated that when subjects injected with adrenalin did not anticipate that arousal effects would result, they tended to attribute their emotional state to the situation, feeling happier or angrier depending on the specific post injection experimental context. That study was the first in a series by Schachter, his students, and others that indicated that people are quite flexible in the manner in which they make causal inferences about the source and meaning of their emotional arousal.

Subsequent research exploring the implications of those findings for moral behavior followed the work of Lykken (1957). Lykken had shown that psychopathic individuals seemed not to experience normal emotion-induced inhibitions. Schachter and Latané (1964) demonstrated that psychopathic criminals (compared to normal criminals) would learn to avoid shocked errors (compared to nonshocked errors) in a lever-maze task only after sustaining arousal from an adrenalin injection. Schachter and Ono (cited in Schachter & Latané, 1964) found increased cheating by college students who had received the tranquilizer chlorpromazine (compared to placebo control groups). These studies can be viewed as evidence of the necessity of emotional arousal as an important component in the avoidance of responses with aversive consequences, with an increase in arousal facilitating that avoidance and a decrease attenuating avoidance. However, since only the degree of arousal was manipulated, and not interferences the subject might make concerning causal attribution, the necessity of arousal was demonstrated but the issue of the sufficiency of arousal was not addressed.

By illustrating the role of emotional arousal, these data provided support for Mowrer’s (1950) two-factor avoidance model. Mowrer’s model emphasized that the emotional response that the organism developed in response to the cues signaling an aversive event would motivate the avoidance response, which, when successful, would lead to the reinforcing reduction of the negative emotional state. Emotional arousal in anticipation of punishment for a potential transgression should, by this interpretation, lead to the avoidance of the temptation behavior. But Schachter and Latané (1964) speculated that for the psychopath, the avoidance of behavior with aversive consequences depended on more than the mere presence of arousal.

More modern theorists have extended this perspective, suggesting that anxiety becomes conditioned specifically to the “behavioral and cognitive precursors of the act” (Aronfreed, 1968, pp. 54–55).
Psychopaths experience a high level of physiological arousal with no apparent subjective emotional distress in response to cues that would evoke an emotional response in normal individuals. Schachter and Latané suggested that it was the chronic inability of the psychopath to experience his physiological arousal as emotion that accounted for his moral insensitivity. These suggestions approach the perspective proposed in this article, but our view differs by suggesting that even in temptation situations where arousal is experienced as emotion, specific causal attributions are crucial in determining the behavioral outcome.

It was the Schachter and Latané (1964) research that led to a series of studies by the first author indicating that college-age men and women more often succumbed to the temptation to cheat when they could attribute the emotional arousal associated with cheating to a placebo pill, which supposedly had associated emotional side effects (Dienstbier, 1972; Dienstbier & Munter, 1971). Whereas Schachter and Singer (1962) had demonstrated that external cues are important in allowing the attribution of and experiencing of artificially induced arousal as specific emotion, the cheating research demonstrated the reverse—that naturally induced specific emotions could be interpreted as nonspecific arousal when subjects attributed the associated arousal to a nonemotionally relevant source such as the placebo pill. Since these cheating studies are important for the theoretical perspective developed in this article, they are reviewed here in some detail.

Recruited as participants in a study on the effects of a vitamin supplement on vision, subjects were given a vocabulary test as a “delay task while waiting for the pill to take effect.” Besides anticipating visual effects from the pill, a random half of the subjects were led to anticipate arousal effects of “a pounding heart, hand tremor, sweaty palms, a warm or flushed face, and a tight or sinking feeling in the stomach” (following Nisbett & Schachter, 1966). Control subjects anticipated benign side effects irrelevant to arousal. After the vocabulary test, to validate the potency of the pill, subjects viewed the auto-kinetic illusion, and were led to understand that the pill was responsible for their seeing the immobile light as moving. Thereafter, subjects were given the opportunity to check over their vocabulary test answers and to see the correct answers. They were informed that the department developing the vocabulary test was interested in using it for future prediction of college success, and that the “board of psychologists” responsible for it would like to interview any students who did particularly poorly on the test (less than 20 right). The test was designed so that less than 1% of college freshmen could attain that score without cheating.

The period during which the individual could look over the correct answers to the vocabulary test and his own answers was also introduced as a “delay period,” in this case prior to the rating of side effects because “it was important that the subjects rate the side effects they might be experiencing just 9 minutes after they had experienced the visual effects.” The salience of the side effects expectation was increased for that period by having subjects look over a side effects checklist at the onset of that period. Any cheating that took place during the delay period could be detected because pressure-sensitive paper (making a second copy of the answers) was placed under the answer paper during the taking of the test; the answer paper was removed by the subjects prior to the cheating period, so that thereafter any changes made on the real answer paper would not correspond to those on the pressure-sensitive paper.

2 Suggestions by Mandler (1964) and the theoretical structure proposed by Lazarus (1968) are useful in understanding possible mechanisms to account for the psychopath’s lack of emotional experience in the face of arousal. Mandler suggested that the childhood of the typical psychopath is characterized by frequent emotional upheaval— and that the individual might learn to minimize the emotional discomfort of these frequent traumas by ignoring the physiological signals from his body, which, if allowed to contribute to a strong emotional feeling, would greatly accentuate his distress.

Mandler’s suggestions fit well with Lazarus’s (1968) concept of benign reappraisal, the process by which the stress reaction is reduced by a cognitive reappraisal involving a lowered assessment of the potential danger in a situation after the individual has adjudged that he is powerless to affect the outcome.
In the first study (Dienstbier & Munter, 1971), 49% of the subjects who anticipated arousal side effects from their placebo pills cheated, compared to 27% of the control condition, with cheating defined as changing any answers. This significant difference became more impressive when it was noted that the entire effect was due to the male subjects, whose cheating rates were 56% and 17% in the two conditions, respectively. Follow-up research (Dienstbier, 1972) indicated that women, too, respond to the arousal placebo with increased cheating if the threat of the consequences for failure is made less frightening (it appeared that in the original study the more intense anxiety experienced by women in response to the threat of being called by the “board of psychologists” interfered with their attending to the pill side effects). In a second study reported in the same article, it was further determined that the effect of the arousal placebo in reducing resistance to temptation was indeed due to the subject’s attributing specifically to the pill—merely attending to arousal feelings without having taken the placebo did not produce the effect.

It appeared that all subjects experienced emotional arousal while confronting the temptation to cheat, largely as a result of considering the act of cheating and the potential negative consequences. But when subjects could attribute the cause of their arousal to the pill, that arousal ceased to have an inhibiting effect. It is not, in other words, that the subjects in either pill condition experienced more or less arousal, but rather that the attribution of the arousal to a neutral source (in the arousal pill condition) defused that arousal, preventing it from having emotional relevance.³

³ There are two lines of evidence suggesting that subjects make the attributions of emotional causality at a less than verbally conscious level. First, when subjects were interviewed, none ever indicated conscious awareness of such a decisional process. (Subjects were completely debriefed and questioned concerning possible suspiciousness following the cheating research.) Second, at the end of the cheating period, when arousal subjects rated the degree to which they experienced the arousal side effects, the rate of reporting such side effects was surprisingly low. (There was, unfortunately, no control group with which they could be compared on this dimension, since the benign symptom subjects rated themselves on the benign symptoms.) It appears, therefore, that although the information concerning expected symptoms is given verbally, thereafter the very significant impact that these symptom expectations have upon cheating behavior is achieved by processes not apparent to the subject.

Whereas Schachter and Latané (1964) had speculated that this process in chronic form accounted for the moral insensitivity of the psychopath, the cheating studies suggested that the same process could be demonstrated to produce episodes of moral insensitivity in normal subjects.

The work of Solomon and Wynne (1954) and subsequent researchers emphasized that although emotional arousal was necessary in the acquisition of avoidance responses, the elimination of overt signs of arousal was often observed once such responses had been well learned. Except for the Schachter and Ono data (cited in Schachter & Latané 1964), which demonstrated increased cheating with chlorpromazine, we would not have anticipated, therefore, that emotional arousal played any role in moral decision making for normal adults. Yet the cheating results suggested that even for 19-year-olds in temptation situations, the avoidance of transgression is still heavily dependent on emotional arousal. Additionally, the studies demonstrated that although 19-year-old adults should have had considerable experience in learning which features of their own behavior and of the environment lead to negative emotional arousal, they were surprisingly susceptible to external influences in attributing arousal experienced in the face of temptation.

But however powerful these data are, they do not provide the theoretical bridge to our final point. Our thesis requires that we demonstrate more than the defusing of emotional arousal through attribution to a non-emotional source in a context as artificial as that associated with the placebo pill studies. We must demonstrate instead that emotional arousal may be defused...
for moral behavior even though still experienced as negative emotion, through the attribution to an emotion-inducing source that has no direct relevance to the moral decision.

**Socialization and Emotion Attribution**

We speculated that children would be even more receptive than adults had proven to be to cues that would suggest different causal attributions for emotional arousal, and that these cues would play an important role in determining the impact of socialization processes. Specifically, if a socialization pattern left the child with the impression that the cause of his emotional arousal or discomfort was his own misbehavior, the impact of this internal orientation on the child’s behavior would be very different from the impact of an external orientation in which equally intense emotion was thought to be caused by fear of punishment. The latter attribution would be particularly ineffective in situations in which the tempted individual could be certain that detection was impossible.

Even early in their socialization history, most normal children will experience some emotional arousal following transgression and detection by a disapproving adult. Several features of the situation naturally lead to that arousal: the child’s knowledge that the adult disapproves, previous scoldings now recalled, anxiety over the uncertainty of the outcome, and the likelihood that the adult will emphasize the discrepancy between the child’s behavior and higher standards. The manner in which the socializing agent responds to the child’s transgression will influence the causal attributions the child will make about his emotional discomfort. Socialization procedures that draw the child’s attention to the transgression rather than to the aftermath (confrontation and punishment) facilitate the attribution of arousal to the act of transgression. Generally, such procedures involve discussions of the transgression in a manner that does not threaten the child, so that the child is able to dwell on the transgression itself and the behavioral standards he has violated, rather than on the threat that confrontation offers. Often the self-concept is manipulated (e.g., “you are a bad boy”) in relation to the standard or behavior (“because you didn’t listen and lied”). The attribution of arousal to the anticipated transgression later leads the child to avoid the transgression itself. On the other hand, responses by socializing agents that threaten the child or that become salient through pain draw sufficient attention away from the transgression so that the child is likely to attribute his emotional discomfort exclusively to confrontation with the socializing agent. When later facing a similar temptation situation, a child treated in the latter manner might experience a high level of emotional arousal, but by attributing his arousal to fear of detection and punishment he would tend to resist temptation only if he believed that detection was likely.

In our research with children we wished to buttress our theory by establishing a temptation and transgression situation in which the emotional arousal following detection would be equal for the two groups of children, with the children assured that future transgressions could not be detected. We then imparted to the children an internal (due to own behavior) or an external (due to confrontation) causal attribution concerning the source of that emotional arousal. It was hypothesized that the children who believed that they “felt bad” for their previous transgression (internal orientation, or guilt) would misbehave far less than children in the same situation who believed that they “felt bad” because their previous transgression had been detected (external orientation, or shame).

**STUDY 1**

*Method*

In Study 1 of the two studies with children, second-grade subjects were brought, one at a time, to a “new toy room” in their school and were asked to keep watching a slot car go slowly around an oval track during the experimenter’s absence; they were to stop the car with a control device if the car began to go too fast. Although the child was given sufficient justification for watching the car, other electrical toys “with loose wires” controlled by a hidden observer were started and stopped according to a schedule to distract the child. If the child transgressed by allowing himself to be distracted from
the implied duty of watching the slot car, the observer caused the slot car to jump the track (the event that was supposed to be avoided) and simultaneously signaled the experimenter to return.

Pretending that the slot car was broken, the experimenter asked the child to accompany him to a separate room, where the independent manipulation was administered: labeling the child’s feelings as due to transgression (guilt) or due to being found out following the transgression (shame). The child was then returned to the room, taught to operate the slot car correctly, and assured that this “new” car would not break if it jumped the track; he was also assured that the door would now be locked and the experimenter did not have a key, implying that detection of transgression of failing to watch the slot car would be impossible. The dependent measure was the total time spent not attending to the slot car during this second watching phase of 12 minutes (with the experimenter absent). A reliability of .99 was obtained between two observers for five subjects; thereafter one observer was used.

One male and one female experimenter participated, each using boys and girls as subjects, so that with the independent manipulation of guilt versus shame, a $2 \times 2 \times 2$ factorial design was effected.

Results

Results indicated no significant effects for either the sex of experimenter or children, but the nonsignificant difference between the mean seconds transgressing for the shame condition (164) and the guilt condition (102) were suggestive. A further analysis indicated that most of the demonstrated increased effectiveness for the guilt manipulation was due to the female experimenter—187 seconds transgressing in the shame condition, compared to 80 seconds in the guilt condition, $t(22) = 1.91, p < .07$, two tailed. It was our impression that a real difference did exist, despite the lack of a significant main effect for the shame versus guilt conditions or an experimenter by condition interaction. It was suspected that the differences between experimenters were due to their different experience levels with children (the more effective one was an elementary school teacher and a parent) rather than a true sex difference.

The inconclusiveness of the findings of Study 1 led to the changes incorporated into the second study, which will be presented in somewhat greater detail.

STUDY 2

In an effort to reduce the variations in background between children in the two conditions and to allow a matched-pair type of analysis, second-grade twins were recruited by seeking the cooperation of their parents and of the twins themselves. Except for some minor changes in procedure and script and the participation of only one experimenter (the first author, who fit the criterion of having sufficient experience with children), the design of the study was essentially the same as for Study 1. Children were not considered separately by sex, since the results of Study 1 suggested no sex differences.

Method

The first twin to participate was chosen randomly and brought to the toy room while the other twin played quietly with an assistant in the waiting room. The participating twin was seated in front of the booth and was told that the slot car presently on the track was very old so it might break if it did fall off the track, but that the problem with the cars sometimes going too fast was in the track. The experimenter’s first absence was so that he could “go to call the man who fixes the toys.” The scheduled use of distraction toys such as the electric train “with the loose wires” was essentially the same as in the first phase of Study 1. Although the criterion time before the slot car would jump the track and the booth light would go on for the first twin was 6 continuous seconds of distraction, if the child had not met that criterion by the end of the 6th minute, the criterion time was reduced by 1 second for each passing minute of attentiveness of the child. Since the members of each twin pair tended to act remarkably alike, their first phase participation times tended to be very similar.

Immediately following the child’s transgression, the experimenter returned. Upon “discovering the transgression and the “broken” slot car, the experimenter asked the child to accompany him to the adjoining room, where he identified the twin’s experimental condition and administered the appropriate manipulation according to the following script (the sections in parentheses indicate the additions that change the basic guilt script to

4 A role filled by the second and fifth authors.
5 A role filled by the third and fourth authors.
the shame script, or where two sets of parentheses occur together, the first is the shame section and the second the guilt section:

I bet you feel a little bad now that (I found out) the car fell off, I’ve seen other kids feel bad when (someone found out) they weren’t able to do exactly what they were supposed to do. When the other kids who tried to watch the car couldn’t, they felt bad (when I found out) too.

But it’s important that we do stop the car when the light first comes on, and before it flies off the track, so we’ll try again with this new car. You won’t feel bad if (I find this car still) (this car stays) on the track—if (I find) you’ve done a good job, you’ll feel good, won’t you?

Have you noticed that when other kids (can show people that they) have done the right thing, that they feel good? And when (people find out) they did things they were not supposed to, they feel bad, don’t they? (when someone finds out) (Even if no one ever finds out) they feel bad for not doing what they were supposed to do, don’t they?

After returning to the toy room, the child was directly informed that since he now was being taught to start and control the new unbreakable car, and since the door would be locked and the experimenter had no key, that “even if it did fly off the track I’d never find out, since you could put it back by yourself and I can’t come right in with the door locked.” Most children spontaneously acknowledged that this was true. With that statement, the experimenter signaled the observer to put his hand over his ears while the independent manipulation was administered: “Remember, you’d feel bad if (I ever found out you did a bad job) (you did a bad job, even if I never found out). There is no way that I could find out this time, though.” The experimenter then left the room, locking the door and very loudly checking it to make sure it was adequately locked.

A 12-minute period, spiced with the same distracting toys starting up and the same dependent measure (the time spent not attending to the slot cars) was identical to the second phase of Study 1.

Results and Conclusions

Of the nine pairs of twins who came to the laboratory, complete data were gathered for six pairs. Two pairs were lost because one of the twins failed to follow instructions by shutting off the slot car so that he would not have to watch it, One pair was not run because the twins were hyperactive and inattentive. For each of the six pairs run, the twin receiving the shame instructions transgressed almost twice as much as the twin in the guilt condition. The average time transgressing was 322 seconds for the shame condition, and 177 seconds for the guilt condition \( t(5) = 8.0, p < .001 \), two tailed. The regularity of the effect across pairs of twins is illustrated by Figure 1.

It is apparent that 8-year-old children are sufficiently cognitively sophisticated so that the differences between the guilt and shame manipulations are perceived, even in the face of the anxiety elicited by the confrontation with a mildly disapproving and strange adult. Additionally, the findings of 60% (Study 1) and 80% (Study 2) increases in transgression time for shame subjects over guilt subjects support conclusions from the adult cheating studies—the attribution of emotional causality has a large impact on the behavior of children in resistance-to-temptation situations.

\(^8\) Thanks are due Steve Slane for his faithful and competent filling of the observer role.
The degree to which the two child studies are acceptable as powerful support of the theoretical development is dependent on several assumptions that cannot be empirically demonstrated. Although we endeavored to maintain equal levels of emotional arousal for the children in the two conditions, and although there are no apparent reasons why the shame or guilt script variations might have induced different arousal levels, no direct assessment of emotional arousal was made. Also, we cannot be certain that the greater effectiveness of the guilt script was not due to one of the scripts drawing the subject’s attention to considerations irrelevant to the attribution of emotional arousal. It might have appeared, for example, that the experimenter was less interested in the behavior of the children in the shame condition. Although the similarity of the scripts was meant to preclude this possibility, it is impossible to be certain that the script differences did not trigger different reactions between the children in the two conditions. However, since neither of these two explanations could easily be used to explain the results of the cheating studies with adults, the hypothesis involving the causal attributions of emotional arousal seems the only likely explanation common to both research approaches. Although the following discussion assumes that the child studies provide strong empirical support for the theoretical developments, we believe that it is possible to follow and accept the theoretical aspects of our thesis without being totally convinced of the relevance of the data from the child studies.

One possible approach to explaining the high dependence of our child and adult subjects on external cues rather than past experience in making causal attributions of emotional arousal is to suggest that this flexibility is greatest when the emotional state in question is an uncomfortable one. Some support-by-default for this view comes from the emotion-attribution literature; the impact of causal attribution manipulations has been convincingly demonstrated only with the negative emotions of anxiety, fear, and startle response (Dienstbier, 1972; Dienstbier & Munter, 1971; Nisbett & Schachter, 1966; Ross, Rodin, & Zimbardo, 1969; Storms & Nisbett, 1970; Zanna & Cooper, 1974). In the first author’s research, several failures to observe the causal reattributions of arousal associated with emotions not related to fear reinforced this view. Further support is provided by the research of Lazarus and his associates (Lazarus, 1968), who have demonstrated a readiness in their subjects to use cognitive information that allows the reduction of stress.

Another approach to understanding the apparent ease with which the children accepted the attribution of arousal to their own behavior or internal causes rather than to detection or external causes is to focus on the readiness of children in this age category to accept such internal attributions. Wolman, Lewis, and King (1971) demonstrated that as children pass from “young” (ages 5–7) through “middle” (ages 8–9) to “old” (ages 10–13) categories, they report that the conditions that stimulate emotions more frequently come from within themselves. Kagan (1971) has summarized a wide range of literature suggesting that the child’s capacity for guilt increases most rapidly between the ages of 4 and 10.

RELEVANCE TO MAJOR THEORETICAL APPROACHES

Social Learning Approaches

Influenced heavily by traditional learning theory on the one hand, and by the psychoanalytic concept of identification on the other, the social learning approach to morality has progressed in the direction of the analyses of those features of the model that facilitate identification and those features of discipline techniques that facilitate resistance to temptation and appropriate affect.

Since the impact of modeling and identification research appears to be waning, we will not attempt an integration of our view with that branch of the social learning literature.7

7 The modeling literature, typified by the work of Bandura and his associates (Bandura, 1968) and Sears and his associates (Sears, Maccoby, & Levin, 1957; Sears, Rau, & Alpert, 1965), has never freed itself from the issues of identification versus simple imitation, and in the 1970s seems to have lost its appeal as a research area. The failure of many major studies in this area to find convincing evidence of causal links be-
Instead, our discussion will be concerned with socialization techniques and the impact of those techniques on the affective and cognitive responses that facilitate resistance to temptation. Several issues in the social learning and the socialization-of-morality literature will be viewed from our theoretical perspective.

Punishment. As a behavior modification technique, and as demonstrated in many laboratory studies, punishment (including physical) is often a viable control technique. That punishment is effective in behavioral control (so that its various parameters bear analysis) is seldom questioned, and those laboratory studies that have compared punished and nonpunished control groups have generally demonstrated the superiority of punishment in behavioral suppression (see Johnston, 1972, for a recent review). But the socialization literature based on naturalistic observation in home and school settings indicates that the use of physical punishment is negatively associated with the internalization of social control (Bacon, Child, & Barry, 1963; MacKinnon, 1938; Maurer, 1974; Nowicki & Segal, 1974; Sears, Maccoby, & Levin, 1957; Sears, Rau, & Albert, 1965).

The dilemma of physical or power-oriented punishment being successful in the laboratory and counterproductive in less controlled settings, particularly as compared to induction, relates to our approach according to the following analysis. When all other things are equal, the more severe the punishment, the more the emotional arousal and the more the resistance to deviation in the future, unless the more intense punishment causes different attributions about the source of the emotional discomfort associated with that situation or distracts from other administered messages relevant to emotional attribution. The first part of that statement is an affirmation of classical avoidance theory—that greater arousal will lead to more resistance to deviation. But, as is also suggested, intense physical punishment should indeed cause different attributions of arousal from either mild punishment or other disciplinary techniques that do not direct attention away from the transgression. If the physically punished individual faces future temptation in situations in which detection seems unlikely, then we would predict a paradoxical effect, with more intense punishment or more physically oriented punishment leading to less resistance to temptation.

Punishment that is love oriented, however, seems less destructive in its impact on the development of resistance to temptation. There are several reasons why such techniques are less counterproductive than physical discipline and why they are found to be positive in some studies. While physical punishment is usually accompanied by parental expressions of anger, encouraging similar expressions in the child, love-withdrawal techniques are associated with anger inhibition (Hoffman, 1970) in the child: the experiencing of anger by the child suggests that emotional arousal is being attributed to confrontation with the parent. Additionally, the love-withdrawal techniques tend to leave the child in a state of emotional discomfort for longer periods, with the termination of such punishment often dependent on some action or statement by the child that relates to the child’s initial transgression. Expressions such as “I’m sorry that I did that” are emotional attribution statements frequently encouraged by parents in psychological punishment situations prior to the termination of the punishment. That love-oriented techniques have sometimes been found to be most effective when practiced by warm parents (i.e., Sears et al., 1957) may be due to the child’s finding it more difficult to attribute his discomfort to the warm parents’ behavior, rather than to his own. The fact that with greater warmth, there is more love to withdraw in the punishment situation is also relevant.

Due to the unfamiliarity and associated uncertainty that the young subject experiences in the laboratory, combined with (usually, in our view) insuf-
sufficient assurances of the undetectability of deviance, the laboratory provides neither a setting in which the exclusive attribution of arousal to punishment is likely nor one in which feelings of total freedom from detection are likely. In laboratory studies the punished child therefore usually appears more self-controlled than the nonpunished control. Unfortunately, a complete resolution of this point is not possible by reference to the morality literature, since researchers seldom attempt to provide evidence that their children were convinced that their transgressions could not be detected. Given that the laboratory is often quite unfamiliar to the research subjects, for young children to be secure in their belief of undetectability considerable effort should be devoted to the staging of experimental situations; if possible, the impossibility of detection should be articulated to the child (as in Study 2).

Naturalistic settings often permit the child to make causal attributions that negate the effect of punishment and lead to freedom from detection. In a home or classroom the harshly or frequently punished child has repeated opportunities to make accurate attributions that the negative emotions he experiences are due to confrontation and punishment rather than transgression and to become certain about when detection is possible. Those accurate emotional attributions facilitated by repeated experience in natural settings are not conducive to internalization. With the development of accurate attributions about the source of negative emotional arousal, either of two types of results may obtain (probably most such situations result in some of each of these two processes). The first process has been emphasized by those social learning theorists who have been heavily influenced by the traditional avoidance model exemplified by Mowrer’s (1950) work, but who must explain the data that indicate the ineffectiveness of punishment in natural settings. Simply stated, that view suggests that the punished child learns accurate discriminatory cues so that no arousal results during transgressions that occur in the absence of punishing agents. While it is apparent that such a process may take place, we would argue that there are a great many temptation situations for normal individuals in which arousal is elicited even though the individual is fully aware that detection is impossible; both the cheating studies and our child studies provided situations designed to make detection appear impossible, yet arousal and the causal attributions made about that arousal played a key role in determining the outcome.

We seek what we perceive to be a necessary middle ground between older theories which emphasize automatic conditioning procedures to the exclusion of cognitive processes, and more modern approaches, which emphasize total cognitive control over the learning of contingencies. The older view (Mowrer, 1950), suggesting a classical-conditioning-like mechanism to account for the attachment of negative emotions to situations and behavior, could not account for data such as that offered in the cheating studies or the child studies, in which cognitive interpretations given to subjects relevant to the meaning of their emotional arousal had important effects. More modern approaches, such as that of Bindra (1974), suggest that reinforcing or punishing events do not directly affect the strength of the preceding response; rather what is learned are the “contingent relations” existing between the events of the situation. Bandura (1969, p. 553) suggests a similar relationship in stating that the classical-conditioning-like aversive techniques used in behavior modification procedures affect the client only to the extent that he is consciously willing to recall and reexperience the therapy situation when faced with temptation in later outside situations.

Our view suggests a greater degree of autonomy between cognition and emotion, with cognition influenced by, but not under the total control of emotional arousal, and emotional arousal likewise influenced by, but not completely controlled by cognition interpretations. Thus the cognitive knowledge possessed by particularly the very young child that his emotional upset in the face of temptation stems directly from previous punishment may not eliminate that emotion when detection is thought to be impos-
sible; rather, such a causal attribution will make the emotion irrelevant and will allow the child to act somewhat as if the emotion were not present. To illustrate this cognition-emotion interaction, consider behavior in the face of a mild phobia. The fear experienced by an acrophobic adult while climbing an open stairway may be as intense when the fear is cognitively perceived to be unjustified (the stairs are safe) as when it is perceived to be justified (the stairs may break); yet through moderate fear levels, such cognitions will determine whether the individual will override his emotional arousal and continue to climb. We present the acrophobia illustration as more than a metaphor and suggest that many of the more important moral prohibitions learned by individuals with well-developed consciences have a phobia-like character, with emotional arousal elicited even though the tempted individual may cognitively appraise the situation as perfectly safe from detection. Such an appraisal allows the emotion to be overridden without being eliminated if the individual attributes his arousal to a fear of detection or to other causes that may be irrelevant if detection is impossible.

Recent research by Loftis and Ross (1974) suggests the possibility of a second effect resulting from different causal attributions. After the acquisition of an emotional response, the extinction rate for that response may depend in part on the source to which the individual retrospectively attributes his emotional arousal. In their research, a tone was heard while a shock-induced emotional reaction, as measured by galvanic skin response (GSR), was conditioned to a light. When later instructions indicated that the emotional arousal could be attributed to the tone, extinction of the GSR was faster (compared to a control group for which the tone was not described as arousing), even though no tone was used during extinction. Thus, although our view emphasizes that emotional arousal will have differential effects depending on the causal attributions made, we do not exclude the process of different attributions eventually affecting arousal level. With respect to moral behavior, the two processes should tend to influence responses similarly, with attributions that arousal is due to irrelevant sources (such as fear of punishment when detection is not possible) both negating the effect of the arousal and, with repeated experience, leading to a reduction in emotional arousal as well.

Techniques such as induction (Hoffman, 1970), which do not depend on punishment, will undoubtedly result in less negative emotional arousal associated with transgression than techniques based on punishment; but successful induction, involving the explanation to the child of why certain behaviors are right or wrong, focuses attention on the behavior per se, so that the arousal that is present during such socialization episodes will be attributed to the act of transgression rather than to the threat of punishment. Hoffman’s review indicates that induction is more effective in moral socialization than punishment-based techniques, even including love-oriented punishment. Our emphasis on the necessity of arousal for induction to be effective is relevant to Aronfreed’s (1968, p. 9) suggestion that “available evidence usually points to great discrepancies between children’s verbal expression of evaluative standards and their actual behavior in a real social context.” The data of Henshel (1971), indicating increasing correlations (from fourth to seventh grade) between honesty assessed verbally and resistance to temptation in a cheating task, provide the basis for an explanation of the discontinuity between verbal expression and moral behavior noted by Aronfreed. Henshel suggested that behavior is not dependent as much on simply “knowing” what is right as it is on “feeling what is desirable.” Unless standards become associated with emotional responses, children will be as free to violate these standards as were the

---

8 An extreme example of an apparently similar process is found in the work of Fenz and Epstein (Epstein, 1967) with skydivers. Although emotional arousal as indicated by physiological monitoring remains high as the skydiver gains experience, his view of the sport as not posing objective danger allows the enthusiast to jump. The changing attributions fostered by experience allow the experience of intense excitement where previously fear had dominated.
subjects in the cheating studies who could attribute their arousal to a pill.

Guilt and resistance to temptation. A second dilemma from the social learning literature is that although a correspondence is assumed (in most versions of social learning theory) between guilt (usually as measured by projective techniques) and resistance to temptation, some studies have failed to show such a correspondence (Allinsmith, 1960; Burton, Maccoby, & Allinsmith, 1961; Maccoby, 1959).

Some authors (e.g., Hoffman, 1970) have suggested that one of the major problems with that literature is that researchers have seldom endeavored to measure guilt in relation to the same types of transgression on which the behavioral measures are taken. Furthermore, the difficulty of interpreting projective measures of guilt is often cited. Our approach lends support to that explanation and allows a further articulation of the importance of measuring guilt and resistance to temptation with respect to the same transgression. As demonstrated by the cheating and child research, our approach emphasizes the high degree of dependency on immediate situational cues in attributing causation of emotional arousal. Thus it is quite likely that individuals who make causal attributions about the internal origins of arousal (guilt) in one situation will find that either their own past experience or current situational cues lead them to make more external attributions concerning emotional causality (shame or fear of punishment) in other temptation situations. Significant correlations between guilt and resistance to temptation should, therefore, be expected only when those dimensions are compared with respect to the same behavior. The only study that measured guilt and resistance to temptation with respect to the same type of transgression found that a significant relationship did exist (MacKinnon, 1938).

Cognitive Dissonance and Self-Perception Theory

The changing of the value of certain behaviors and objects through unnecessarily strong negative external inducement (threat) has been demonstrated by researchers investigating a hypothesis derived from cognitive dissonance theory. That hypothesis stated that children who receive a mild threat to inhibit an act will take more personal responsibility for resisting temptation than will severely threatened children; they will therefore experience more cognitive dissonance over resisting temptation. Aronson and Carlsmith (1963) and Freedman (1965) demonstrated that when children were given a mild threat to not play with a valued toy, they valued the forbidden toy less (Aronson & Carlsmith) and played with it less (Freedman) than was the case under severe threat. The sophisticated control conditions of the Freedman study (finding no threat-effect differences when there was no opportunity to play with the toy during the threat-relevant period) allowed the elimination of many potential explanations alternative to the dissonance-derived hypothesis.

A more recent literature has developed indicating similar effects for unnecessarily strong positive inducements (reinforcement). The thinking of deCharms (1968) on the importance of perceived locus of control in motivation led to research on the overjustification effect by Deci (1971,1972), Kruglanski, Alon, and Lewis (1972), and Lepper, Green, and Nisbett (1973). This research indicated that when an individual is offered reinforcement for a task he would have performed without payment, he acts (in future situations without payment) as if he were less intrinsically motivated than before the unnecessary payment.

Both the dissonance and the overjustification research provide support for Bem’s (1972) self-perception theory—that individuals tend to consider all available evidence in assessing the motivation underlying their own behavior, just as an observer looks for all the forces relevant to another’s behavior in making attributions about the other’s motivation. Our approach suggests that the overjustification and dissonance results are due to the effect that an offer of tangible reinforcement or threat of severe and tangible punishment has on reducing attributions about
the internal cause of positive and negative emotions. But by focusing as it does on the ability to change causal attributions of emotion associated with and resulting from behavior after the behavior has been performed, our approach suggests that reinforcement or punishment given after the fact will often create effects similar to inducements and threats made before the fact. This conclusion from our studies with children is reinforced by the findings on “retrospective misattribution” by Loftis and Ross (1974) cited earlier. The importance of this extension becomes more obvious when we consider that socialization efforts are usually made after some motivational failure has occurred.

As suggested earlier, the extent to which punishment or reinforcement draws attention to itself rather than to the preceding behavior will be influenced not only by the specific form of punishment or reinforcement but also by the style of administration and the situation. Thus while tangible reinforcers, such as physical punishment, are likely to draw the attribution of emotional response toward themselves rather than to the preceding behavior, certain approaches may accentuate that process. Information that suggests that the administered reinforcement should be sufficient compensation and an adequate inducement for future performance should increase the likelihood that pleasure derived from the task itself (from the means, as well as from the completion) will be attributed instead to the reinforcement.

This approach may be useful in illuminating an apparent contradiction in the recent overjustification literature. Unexpected prizes administered to a group of game-winning students by Kruglanski et al. (1972) had the effect of reducing liking for those competitive games (compared to nonrewarded winners), while children who received unexpected prizes for good drawing in the Lepper et al. (1973) study did not later behave as if their prizes caused them to devalue the special drawing task used. From our perspective, the critical difference lies in the changed attributions resulting from Kruglanski et al. telling their subjects that they had been promised the rewards prior to the games (which in fact was not true) and in giving tangible reinforcers (plastic puzzles), which were salient during the game ratings because they were given before the rating. The positive emotional response resulting from playing and winning the games was apparently partially attributed to the prizes. In the Lepper et al. study, however, the unexpected reward group was not misled to believe that their rewards had been previously promised nor were those rewards salient at the time the dependent measure was taken; their rewards had been posted (several days previously) on an “award board” that was not present in the classroom in which the dependent measure was taken.

Our approach leads to a theoretical integration between findings associated with achievement motivation and the overjustification research, an integration very similar to that which is apparent between the punishment literature (in natural settings) and the dissonance literature. Research in the achievement area indicates that while gestures of parental love and affection (nontangible reinforcers that do not draw attention from the child’s behavior) are related to high achievement motivation in boys, the administration of tangible reinforcers by parents is not associated with high achievement motivation (Rosen & D’Andrade, 1959; Spence, 1970; Winterbottom, cited in McClelland, Atkinson, Clark & Lowell, 1953).

The Cognitive and Avoidance Theory Interface

Cognitive development theories of morality, as exemplified by the writing of Piaget and Kohlberg, are too extensive to be reviewed here in detail. Rather, cognitive developmental theory will be briefly outlined and the relevance of that approach to the proposed theory will be explicated at a general level.

Piaget (1932) postulated that the development of the child’s moral judgments corresponds to the development of other more general cognitive capacities. Prior to the development of sufficient cognitive flexibility, the child makes moral judgments in strictest conformity to established rules, view-
ing behavior as totally right or wrong, basing that assessment on the outcome of an act regardless of the intent of the actor, and believing in the concept of “immanent justice”—that punishment will automatically and surely follow transgression, with maturity and cognitive flexibility, moral judgments become more flexible, with rules seen as changeable depending on other constraints of the situation; the outcome of a specific act has less impact on the assignment of blame than the intent of the actor. In part, this transition to flexibility is made possible through active participation in society and through rule making and changing, allowing rules to be seen as serving man’s needs. Social participation also encourages the acquisition of increased empathy through interaction with peers.

Kohlberg (1958) extended Piaget’s approach to a 6-stage sequence, each stage of which is defined by a different basis for moral judgment. In earliest stages, avoidance of punishment (Stage 1) and later the definition of right and wrong based on the valence of the outcome for the actor (Stage 2) prevail. Middle stages involve judgments based on the maintenance of social approval (Stage 3) and then on duty to laws and the prevailing social order (Stage 4). Higher stages are based on an understanding of the social-contract nature of social and legal systems and respect for the rights of others (Stage 5) and finally a recognition that legal and moral principles are often in conflict, resolved through principles of conscience, logical consistency, and mutual trust in fulfilling higher values (Stage 6).

Neither Piaget nor Kohlberg has extensively addressed theoretical questions of the relationship of moral behavior to moral reasoning, or of the relationship of affect and emotional arousal to moral reasoning. Kohlberg (1969), however, approached our viewpoint by suggesting that affective responses would progress developmentally with cognitive responses and that for the higher moral stages, guilt would be associated with different types of transgression. Thus for a Stage 6 individual, “guilt over violation of internal principles” would result; for the Stage 4 individual, guilt would imply “concern about one’s responsibility, according to rules”; for the Stage 3 individual, only shame would be experienced.

Kohlberg’s astute analysis is completely compatible with our orientation, but it is through the emotion-attribution perspective that we can address the issue of the dynamics that account for this presumed usual compatibility of affect and moral judgment in the maturing individual. The key to our explanation is that temptation situations are basically unchanging as the individual matures through the cognitive stages postulated by the cognitive-developmental theorists. That is, even though perceived somewhat differently with maturity, and even though different responses become available, the basic issues that lead to the emotional arousal associated with temptation remain essentially the same. It is, of course, impossible to prove this point, but it is difficult to name any moral issue faced by a mature adult in any sphere of personal or professional activity that does not have an obvious and basic counterpart for a 6-year-old child. There are mature and immature forms of stealing, of being dishonest or disloyal, of hurting others or being unfair, or of being destructive to property and life. This view is supported by the moral dilemmas used in the Kohlberg (1958) system to assess maturity of moral judgment—the situations can be seen as posing ethical dilemmas to individuals at any stage of moral maturity.

How does this continuity of moral dilemmas provide a mechanism for understanding the (normally) continuing compatibility of moral judgment and its supportive affect as the individual’s level of moral maturity changes? A basic tenet of our approach is that many moral dilemmas evoke a phobia-like response, causing emotional discomfort as the individual begins to seriously consider transgression as an option. But that emotional arousal will have different meanings (and will affect behavior differently) for individuals at different levels of moral development. Thus one’s stage of moral development does not determine whether an individual will experience arousal, but rather once such arousal is experienced, one’s stage of moral maturity will influ-
ence what the arousal means and therefore whether the arousal is relevant to the available behavioral options. We do not wish to exclude the possibility that the level of maturity will influence one’s perception of the situation and that this perception will influence the amount of arousal experienced, but the impact of such initial perceptions on arousal level may be minimal.

As an example, we might consider the problem of killing others in a war, assuming a society in which parental instruction, the major institutions of society, and the child’s personal encounters with death all create in the child a negatively toned emotional arousal at the prospect of causing another’s death. Other factors that might support this arousal include the recognition of the extraordinary exercise of power in killing, the irreversibility of death, and the degree to which the behavior is relevant to the support of or the negation of self-concept. When ordered to kill in war time, individuals at different stages in Kohlberg’s (1958) moral development scheme would act differently. The Stage 4 individual, believing firmly in the supremacy of the law and the absolute need to uphold the law, would agree to kill, yet he would experience inhibiting negative emotional arousal (based on processes such as those discussed earlier), which cannot be eliminated readily through cognitive control. Unless overwhelmed with negative emotion, his attribution of affect to a fear of killing would be sufficient to make that negative affect less relevant, allowing him to kill (just as the shame-oriented child is able to transgress even when aroused if detection seems impossible, or as cheating is facilitated for the individual who can attribute his arousal to a placebo). With killing experience, the Stage 4 soldier’s emotional response should decline relatively rapidly (since, as indicated by Loftis and Ross, 1974, the attribution of arousal to an irrelevant source may facilitate subsequent extinction).

Individuals at Kohlberg’s Stage 6 would initially experience arousal for the same reasons and in a manner similar to the Stage 4 individual. But principles based on individual conscience allow more latitude of outcome. One may hierarchically arrange personal values so that killing is required by conscience to defend principles more important that the deaths involved, or one may decide that killing cannot be justified by the principles defended by the war. In the first instance, the arousal would be experienced as guilt (the individual assumes personal responsibility for the decision) over the anticipation of killing, but the attribution of that arousal to the violation of moral principles would not prevent the killing—the arousal would be of little relevance since other moral principles associated with the nominal purpose of the war have become more important.

The Stage 6 individual who has concluded that life is more important than those principles defended by the military action, on the other hand, experiencing his arousal as anticipatory guilt over the killing, would focus on his arousal as relevant to that behavior. Therefore, the arousal could be experienced as a manifestation of the degree of opposition the individual feels toward killing and as evidence of the correctness of the decision not to kill. As the three hypothetical individuals approach their battlefield decisions, increasing arousal will therefore have different impacts on each. The Stage 4 individual, feeling his arousal as unwelcome fear, might begin to feel cowardly and ashamed of his emotional reaction. The first Stage 6 individual, feeling anticipatory guilt, would feel more confused, wondering if this increased arousal should be taken as evidence of the wrongness of killing and hence of a need for the reordering of those values that favored killing. The second Stage 6 individual would, of course, become more convinced of the correctness of his decision not to kill.

Emotional arousal and cognition are thus seen to interact in a complex manner, with behavioral outcome determined by the outcome of that interaction. When either is overwhelming, it may tend to dominate despite the counterinfluence of the other, but in the normal middle ground of life, that mechanism which permits us to display our humanity (when humanity is defined in the old-fashioned manner as the dominance of the rational over the appetitive and emotional) seems to depend heavily on our ability to experience our emotion as inoperative by attribut-
ing it to causes not relevant to the present situation. Tomkins (1963) wrote that it is “one of the major functions of any negative affect theory to guide action so that negative affect is not experienced,” and that this represented “affect at a distance” (pp. 320, 321). We would agree, but add that the major function of any theory that seeks to integrate cognition, affect, and behavior must provide an explanation of how cognition, more differentiated and complex than affective responses, may moderate the impact of emotional arousal on behavior, even while failing to completely dominate or eliminate these affective responses.

References


Submitted December 4, 1974.