The paper investigated the relationship between how people reason about moral matters and how they act. Moral reasoning is defined as the degree of maturity of judgment relating to moral matters. Moral behavior is interpreted to include behaving in ways which are helpful to others in need or distress and refraining from behavior which is physically or psychologically harmful to others. Questions investigated included: 1) do people who reason at more advanced levels act differently in moral situations than those who reason at less advanced levels? 2) do people who reason at the same stage behave similarly in moral situations? and 3) do people act in accordance with what they say they will do in moral action situations? Social science literature, and, particularly, social psychological research studies, are reviewed regarding their treatment of these concerns. Findings indicated that there is some support for the notion that the level of moral reasoning is associated with an increase in frequency of moral behavior and that there exists quite a discrepancy between what people think they would do and what they actually do in moral action situations. Additional research is reviewed regarding factors in addition to moral reasoning which influence moral behavior. (Author/DB)
Is there any relationship between how a person reasons about moral matters and how he or she acts? Do people who reason at the so-called "higher" or "more advanced levels, for example, act any differently in moral situations than those who reason at the lower levels? Do people reasoning at the same stage behave similarly in moral situations? Do people behave similarly in such situations to the way they think they would behave? Questions like these, and their implications for social studies education, are the focus of this paper.

Common sense expectations might lead one to think that moral reasoning and moral behavior are positively correlated -- that the "higher" the stage (in Kohlbergian terms) at which a person reasons, the more likely he or she would be to behave morally if given an opportunity to do so. By "behave morally," I mean either to refrain from acting in ways that most people in our culture would agree could physically or psychologically harm others (e.g., striking, verbally abusing, cheating, stealing from, lying to, etc.) or, to the contrary, to help others in need or distress (e.g., by in some way comforting them, providing assistance to them, empathizing with them, etc.). Indeed, Kohlberg (1972)

has argued that moral maturity in judgment and action are closely related, that "... advance in moral judgment seems to correlated with more mature moral action. Principled subjects both cheat much less and resist pressures by authorities to inflict pain on others much more than do less mature subjects" (p. 79). Is this true?

There are five studies which lend support to the notion that level of moral reasoning does contribute to differentiation in amount of moral behavior. In a study conducted at the college level, Brown and colleagues (1969) found that only 11 per cent of subjects at the principled level (stages 5 and 6 on Kohlberg's moral judgment scale) cheated on an examination as compared with 42 per cent at the conventional level (stages 3 and 4). There were no individuals who reasoned at the preconventional level (stages 1 and 2) in this study. In another study, (Krebs and Kohlberg, 1973) 100 sixth-grade children were given experimental cheating tests and a Kohlberg moral judgment interview. Eighty per cent of stage 5 subjects in the study resisted the temptation to cheat compared with 45 per cent of those at stage 4; 22 per cent at stage 3; 36 per cent at stage 2; and 19 per cent at stage 1. There were no individuals reasoning at stage 6 among the subjects of this study. In 1969, Haan and her colleagues (Haan, Smith, and Block, 1969) studied the situation at Berkeley during the late 1960s when many students were faced with deciding whether to "sit-in" at the
Administration building in the name of freedom of political communication. The researchers administered moral judgment interviews to over 200 of these students. They found that 80 per cent of those reasoning at stage 6, and 50 per cent at stage 5 sat in, compared to only 10 per cent of those at stages 3 and 4. During that same year, Kohlberg (1969) administered a moral judgment interview to some of the participants in the Milgram (1963) obedience study. (Under the guise of an experiment, undergraduate students at Yale were ordered by an experimenter to administer a series of increasingly severe shocks to a victim, who was actually a confederate of the experimenter.) Seventy-five per cent of stage 6 subjects quit or refused to shock the victim compared with only 13 per cent of all subjects at stage 4 or below.

Lastly, Krebs and Rosenwald (1973) report a study whose results are perhaps the most supportive of all for the reasoning-action correlation. Subjects were recruited by an advertisement and asked, among a number of things, to fill out and mail back a questionnaire. The female investigator, a university student, paid them in advance for doing this and gave them a self-addressed, stamped envelope. She then stated that unless she received all of the questionnaires back from them within a week, she would most likely fail her research course. But also, she said, she trusted them to get the questionnaires back to her in time. A moral judgment interview was administered to each of the subjects. There were no stage 1
nor stage 6 subjects in the study. All (100 per cent) of the stage 5 subjects, and more than 70 per cent of the stage 4 subjects, returned the questionnaire on time, compared to only about 30 per cent of the stage 2 and stage 3 subjects.

There is some evidence, then, to support the idea that level of moral reasoning attained does increase the likelihood of individuals acting morally. These studies do appear to suggest, if only somewhat, that in general a greater percentage of higher stage reasoners will engage in moral behavior than lower stage reasoners. Nevertheless, their results must be taken with a grain of salt. First of all, the number of individuals identified as reasoning at the higher levels has been quite small (e.g., in the Brown et. al. study, there were only nine subjects at the principled level; in the Kohlberg study, only six). Second, the results obtained in some of the studies are ambiguous. In the Krebs and Kohlberg study, for example, more of the stage 2 subjects (36 per cent) than the stage 3 subjects (22 per cent) resisted the temptation to cheat. In the Haan, et. al. study, as many stage 2 subjects (80 per cent) sat-in at the administration building as did stage 6 subjects. Third, the conditions under which the studies were conducted (often on college campuses) are not the sorts of situations, it might be argued, in which most people are likely to find themselves.

Furthermore, there are some additional studies, conducted by social (as opposed to developmental) psycholo-
gists, which complicate the picture considerably. According to Kohlberg (1973), the most common level of moral reasoning is the conventional (stages 3 and 4). The most common stage for most adults, in all of the societies he has studied, is stage 4, the "law and order" stage, where the right thing to do is that which one's society defines as right, either by law or by established rules. Large numbers of adults also have been found to be at stage 3, the "good boy-good girl" level of morality, where the right thing to do is that which pleases or helps others, and which is approved by others. Yet the work of Asch (1952); Milgram (1974); and Latane' and Darley (1970) suggest that many individuals, regardless of their stage of moral reasoning, behave at times in ways which they think are wrong, which hurt others, or which violate what society in general considers to be right.

Asch (1951, 1952, 1956), for example, in what are now viewed as a classic set of experiments, asked for volunteers to participate in a study on perceptual judgment. There were 123 subjects (mostly college students). They were placed in a room with six to seven students (who, unknown to the subject, were confederates of the experimenter). The experimenter then showed the group a straight line (X), along with three other lines (A, B, C), and asked each member of the group to say out loud which of the three lines was closest in length to line X. The correct answer is B, but all of the confederates say line A. Each subject was asked to make a total of twelve perceptual judgments involving differing
lengths of line, with one of the lines in the group of three always equalling the length of line X. Since the perceptual judgment to be made each time was a very easy one, Asch expected that almost every one of his subjects would report accurately what they saw. And about two-thirds did. But one-third of the subjects did not. What is interesting for our purposes here is that when Asch interviewed the subjects who did not report, their perceptions accurately, almost all stated that they viewed accurate reporting of what they saw as being "honest" or "conscientious;" not one said it was right to go along with the group. There was, for these subjects, no real uncertainty about what was the right thing to do, yet they did not do it. But even among the two-thirds that did not go along with the group, very few engaged in moral reasoning. The great majority came up with imaginative sorts of explanations for the differences in their own and the group's perceptions. Some said they were viewing the lines from a different position than the others; some said it must have been an optical illusion; some said that it was due to the fact they wore eyeglasses; etc. Interestingly, quite a few of this group (the nonconformers) thought it was possible, and even likely, that the group judgment was right and theirs was wrong. Crutchfield (1955, 1959) conducted similar studies (although his subjects were not face-to-face with the confederates) on more than 600 people, all "above average" in intelligence, education, and occupational status, and
observed similar results.

In the Milgram (1963) study mentioned earlier, 40 subjects (Yale students) served as "teachers" of another subject (once again, a confederate of the experimenter). Every time the learner gave an incorrect response to a verbal learning task (naming correctly the second of a pair of previously memorized stimulus words when given the first), the teacher was ordered to administer an increasingly severe electric shock, ranging from 15 volts ("slight shock") to 450 volts ("Danger: Severe Shock") by pushing a button on a "shock generator." The voltages were clearly labeled, as were the written warnings. The learner responded to the shocks with, progressively, grunts, verbal complaints, cries of pain, pleas to be let out of the experiment and, at the higher levels, screams of agony. Twenty-six (about two-thirds) of the subjects obeyed the experimenter's commands fully and administered the most severe shock (450 volts) to the learners. Fourteen subjects broke off the experiment at some point after the victim protested, and refused to participate any further.

What is interesting with regard to these studies, as in the work of Asch (1952), is the discrepancy which exists between what people think they would do and what they actually do. Milgram (1963) asked other students and colleagues of his at Yale what percentage of subjects they thought would go all the way and administer the most severe shock. They estimated
only 1 per cent. Similar results among university students are reported by Aronson (1972). Forty psychiatrists whom he consulted (Milgram, 1965) were even more optimistic. They estimated only 0.1 per cent. As we have seen, they could not have been more wrong. Two-thirds of the subjects went "all the way." It seems logical to assume that many of the subjects, if they had been asked beforehand, would have made a similar prediction, since some of them were taken from the same population as the judges -- Yale students.

Furthermore, once the subjects learned the true nature of the experiment (all subjects were debriefed once their participation had ended), the obedient subjects almost without exception stated they thought they had done a "wrong" thing, heaved sighs of relief, or shook their heads in apparent regret (Milgram, 1974). Many described their participation as a very painful, even agonizing experience. Most reported undergoing extreme stress and tension (Milgram, 1963). As one observer noted:

I observed a mature and initially poised businessman enter the laboratory smiling and confident. Within 20 minutes he was reduced to a twitching, stuttering wreck, who was rapidly approaching a point of nervous collapse. He constantly pulled on his earlobe, and twisted his hands. At one point, he pushed his fist into his forehead and muttered: "Oh God, let's stop it." And yet he continued to respond to every word of the experimenter, and obeyed to the end (Milgram, 1963, p. 377).

Those who had resisted the orders of the experimenter, on the other hand, felt proud of themselves. It appears that once they were no longer in the experimental situation, the obedi-
ant subjects viewed the situation in the same way as the judges.

In another type of study, Latane' and Darley (1970) and their colleagues were interested in investigating whether bystanders would come to the aid of a victim in apparent distress. In a series of studies, they constructed an imaginary situation [imaginary, that is, to the experimenters] in which subjects perceived an individual suffering an injury or accident and crying out for help. The subjects participating in the study were then observed to see if they would come to the victim's aid. For example, in one study (Latane' and Rodin, 1969), a female experimenter asked a number of subjects [university students] to fill out a questionnaire. While they were doing so, she moved to another room within earshot of the [separated only by a collapsible curtain], informing subjects that she would return when they finished the questionnaire. A few minutes later, she staged an "accident" by playing a tape recording of a young woman climbing a chair, followed by a loud scream and a crash, as if the chair had collapsed. They then heard the woman moaning and crying and saying, "Oh, my God, my foot ... I ... can't move it. Oh ... my ankle ... I can't get this thing off of me." The cries continued for a minute and then there was silence.

An interesting variable in this study was whether or not subjects were alone or with another person as they were filling out the questionnaire. Of those participating in pairs,
some 70 per cent offered to help the young woman; of those participating alone, only 20 per cent offered help. Studies like these (usually referred to as the bystander intervention studies) have two basic characteristics: (1) the need of the person in distress is potentially severe; and (2) bystander intervention is necessary if the distressed person is not to worsen. They have been repeated using many formats, with the bystander subjects being exposed to individuals suffering a variety of mishaps, including an epileptic attack (Darley and Latane', 1968), falling from a ladder (Bickman, 1972), collapsing from a fainting spell (R. Smith, et. al., 1973), groaning and lying in a doorway (Darley and Batson, 1973), or experiencing severe stomach cramps (Staub, 1974). In all of these studies, it was found that bystanders are more likely to intervene when they are by themselves than when they are in pairs or larger groups.

But many do not intervene at all, even when the costs to themselves of doing so are quite slight. Evidence of this is found in another study by Latane' and Darley (1970). In this case, they wanted to investigate whether observers to a theft would report it. They had two "robbers" (actually confederates of the experimenters) enter a discount beer store and ask the cashier at the counter to check how many bottles of Lowenbrau beer he had in stock. While the cashier (also a confederate) was in the stockroom, the two robbers took a case of beer, commenting while they did so "They'll never miss this." They then carried the case to
their car. The robbers bided their time before taking the case so that they could do so with only one onlooker in the store half of the time and two onlookers the other half. The cashier then returned to the counter and resumed waiting on the genuine customer(s). Although they were given ample time to spontaneously report what had happened, only 20 per cent of all subjects did so. When no report was made, the cashier prompted the onlooker(s) by asking what had happened to the two men and if the customer(s) had seen them leave. Putting all reports together, prompted or spontaneous, 65 per cent of the onlookers reported the theft. The percentage of reporting was less, however, when there were two onlookers as compared with one (56 per cent compared to 65 per cent). Thus, even when the nature of the "intervention" is only verbal, and the possibility of retribution is very slight, many people refrain from becoming involved. Why? Is it perhaps that many people just don't care about people in distress or are not interested in what happens to others?

This does not seem to be the case. There is evidence that people are anything but nonchalant about perceiving someone in distress. They are not apathetic or uninterested. Interviews with bystanders to the Kitty Genovese murder in New York City in the late 1960s (38 witnesses watched from their windows for 30 minutes a young woman being murdered) found that all were horrified by what they had seen. (Rosenthal, 1964). But not one even called the police.
How can this be explained?

One thing seems clear. Level of moral reasoning, in and of itself, does not adequately explain the presence or absence of moral behavior in people. Unfortunately, we do not know the stage levels at which the subjects in these studies reasoned, for moral judgment interviews (except for the Kohlberg [1969] study) were not conducted. But it is likely that most reasoned at the conventional level, since it is the most common level attained among adults. Yet, as we have seen, large numbers of the subjects behaved in ways which they later admitted they believed to be wrong, or which hurt people. Furthermore, with regard to those studies in which stage levels were determined (i.e., in the Brown, et al., 1969; Krebs and Kohlberg, 1973, and Kohlberg, 1969, studies), several of the principled level subjects did cheat or administer shocks to innocent victims (11 per cent in the Brown study; 20 per cent in the Krebs and Kohlberg study; 25 per cent in the Kohlberg study). People can fail to act morally, it appears, at any level of moral reasoning, even the highest.

What else, in addition to moral reasoning then, might determine (if only in part) when people will act morally? The Asch ([1952], Milgram [1963], and Latane' and Darley [1970] studies have been replicated a number of times, using many variations in format. These variations have had a considerable effect on the results obtained, and accordingly suggest several factors, in addition to (or perhaps in place of) moral reasoning which help to explain when moral behavior is likely to occur. Asch, for example, found that the fewer the
number of individuals arrayed against the true subject, the
easier the perceptual judgment to be made, and the presence
of even one other "truthful" subject decreased the amount
of conformity by subjects to the views of the group [Asch, 1952].

Variations of the Milgram [1963] experiment, using
over 1000 male and female subjects of all ages and occupa-
tions, showed that the "immediacy" of the victim makes a con-
siderable difference. When subjects were ordered to press
the victim's hand to a shockplate, for example, as opposed
to just pushing a button to administer a shock to a victim
in another room, 70 per cent defied the experimenter. Obedi-
ence also dropped sharply as the experimenter became removed
from the immediate vicinity of the subjects. Thus almost
three times as many subjects obeyed the experimenter when
he was physically present as when he gave his orders by
telephone [Milgram, 1974]. Furthermore, when subjects saw
other subjects refuse to obey the experimenter, 90 per cent
did likewise [Milgram, 1965]. The directness of a subject's
responsibility for administering the shock treatment also
was significant. When they were required only to pull a
master switch to release the actual shock-delivering switch,
but were not ordered actually to deliver the shock themselves,
37 out of 40 adults in New Haven continued to the most severe
level. They said later that the final switch-puller had the
real responsibility [Milgram, 1967].
In the bystander intervention studies, the interesting thing is that an increase in the number of bystanders decreased rather than increased intervention. The most frequent explanation for this is that when others are present, it is easier for an individual to diffuse responsibility, that is, to say in effect that "someone else will help (the victim)." When a solitary individual, on the other hand, perceives someone in trouble, the responsibility for helping that person may not be diffused.

It becomes evident, then, that there are many factors which need to be considered by anyone wishing to make predictions about when, if at all, moral behavior will occur. These factors include the nature of the situation, the number and proximity of other people involved, whether there is support from others to act morally, the nature and source of that support, the kind of decision and/or action required, whether responsibility for action can be diffused to others, and the directness of one's own responsibility for acting morally. Moral reasoning is only one of many possible determinants of moral action. To the extent that the nature and strength of these [and undoubtedly other] factors can be identified in a particular situation, the predictability of moral behavior is likely to be more accurate.

* * * * *

The preceding discussion suggests a number of things for social studies educators to consider, particularly those
who take values education as their particular area of interest and concern, if they wish to contribute to an increase in the likelihood of moral behavior occurring.

First, it must be realized that the development of moral behavior is by no means a simple matter. For one thing, whether a particular kind of behavior can even be considered moral or not depends on how one defines the term "moral." Specific acts without interpretation cannot be evaluated at all. Thus, one problem which confronts the concerned educator in this regard is to decide, and to help students decide, what "acting morally" (i.e., moral behavior) means. What makes a particular way of acting right or wrong, good or bad? And why? What characteristics do moral acts possess which non-moral and amoral acts do not? Trying to decide this question (if only somewhat) is essentially an exercise in concept development, and can be promoted in the classroom using the basic sorts of categorizing strategies that are common in the literature (e.g., see Fraenkel, 1980, pp. 176-87).

Second, once teachers are at least somewhat clear about what acting morally involves, and how moral and non-moral acts differ, they must decide whether they want to try deliberately to foster such behavior in the classroom. This, of course, is a question of value that some say does not fall within the province of the school, but I would argue that to ignore it, or to answer it in the negative, is more
than likely to result in some (quite dissatisfactory) values being taught by default.

Assuming that the answer to the preceding question is in the affirmative, teachers then need to plan how best this can be done. The particular techniques to be used might vary considerably. Although didactic methods have generally been viewed as having little, if any, effect on changing behavior (Hartshorne and May, 1928-30; Festinger, 1964), there is some recent evidence that the intensity with which behavior is condemned has a considerable effect on how young children perceive such behavior (Yarrow, 1977). Although this work has been done only with parents and their very young children, its results merit investigation with teachers and students (of all ages) as well.

The current favorite among many, if not most, educators so far as techniques go is the discussion of moral dilemmas, although no studies have as yet been performed to investigate the effect(s) of such discussions on moral behavior. Evidence does exist, however, to show that it does produce a change upwards in moral reasoning, although not all students advance consistently. Furthermore, such discussions do not appear to be very effective in stimulating reasoning much beyond stage 4 (Lockwood, 1978).

At any rate, far more thought and research needs to be given to what sorts of activities are likely to promote (or at the very least make students aware of) moral behavior. The Asch (1952) studies, for example, suggest that if teachers wish to decrease conformity in students, they might try to
provide more opportunities for students to give support for each other in decision-making situations rather than requiring them to compete against and try to outdo one another. This conceivably might promote a greater amount of self-confidence in students concerning their own viewpoint when they have reason to believe their viewpoint is right, even though it conflicts with the views of others.

An alternative way of conducting moral dilemma discussions is also suggested by Asch's work. The usual format recommended to teachers wishing to conduct such discussions is to have students answer a series of questions designed to get them to think about the issue involved. They then are asked to give their ideas about what they think the protagonist should do, and why. Asch's work suggests, however, that it might be beneficial to ask students to work together, brainstorming as a group (rather than individually) various things which might be done, and of these, which might help the most people involved; in effect, to seek to support, rather than argue against, each other in their reasoning.

The results of Milgram's (1974) work lend support to the idea of encouraging support for others among students. You will recall that when subjects saw other subjects refuse to obey the experimenter, 90 per cent did likewise. Perhaps providing students with many more opportunities to present reports and give demonstrations, to role-play conflict situations, indeed, even to take examinations, which require col-
laboration and mutual support, rather than solitary research and presentation, might increase their ability to identify with, and take into account, the needs of others.

Milgram's work also suggests something else. Again, recall that the directness of the subject's responsibility was a big factor in whether he or she obeyed the experimenter's orders fully. Thus, increasing opportunities for students to be directly responsible for initiating and carrying out individual projects, along with insuring that they experience the consequences of being fully responsible for such (as well as seeing what happens when they are released or absolved of such responsibility) might contribute to a greater incidence of moral behavior.

As no doubt you have perceived, there is the possibility of working at cross-purposes here. The Milgram studies suggest that people are more likely to act morally when they are supported by others in doing so. The bystander intervention studies reveal that solitary individuals are more likely to aid others in distress than are pairs or larger groups of people. The challenge, it would appear, is to engage students in activities which maximize their opportunities to work with and draw support from, yet not be able to diffuse responsibility to, others. It is a challenge to which few, if any, social studies (or other) educators have as yet addressed themselves.
References


Asch, S. E. Studies of independence and conformity. A minority of one against a unanimous majority. PSYCHOLOGICAL MONOGRAPHS, 1956, 70, (9).


Crutchfield, R.S. Conformity and character. AMERICAN PSYCHOLOGIST, 1955, 10, 191-198.

Crutchfield, R.S. Personal and situational factors in conformity to group pressure. ACTA PSYCHOLOGICAL, 1959, 15, 386-388.


Festinger, L. Behavioral support for opinion change. PUBLIC OPINION QUARTERLY, 28, 1964, 404-417.


