

Abstract

Moral discussion is at the heart of moral education, much of which is derived from the cognitive-developmental school of psychology. Much of the literature on moral education is based on the psychological research emanating from that theoretical model. In this paper, the relationship of that research to moral discussion will be examined. A critical review of the research as well as the educational prescriptions derived from it will highlight one particularly controversial aspect of moral discussion, i.e., the “+ 1” convention. Conclusions and revisions will be presented based on the foregoing analysis.

Marvin W. Berkowitz*

A Critical Appraisal of the Educational and Psychological Perspectives on Moral Discussion

The cognitive-developmental model of moral reasoning (Kohlberg, 1976) has been widely applied to education. Such applications have included classroom discussion (Blatt & Kohlberg, 1975; Colby, Kohlberg, Fenton, Speicher-Dubin & Lieberman, 1977; Galbraith & Jones, 1976; Hersh, Paolitto & Reimer, 1979), governance and moral atmosphere (Kohlberg, 1978; Hersh, Paolitto & Reimer, 1979; Power & Reimer, 1978; Wasserman, 1978), and psychology and counseling (Hennessy, 1976, 1979; Kohlberg & Wasserman, 1980; Sprinthall, 1980). All of these applications have two things in common. First, they all are based on the psychological theory and research emanating from the Kohlberg camp. Second, they all include some manifestation of peer moral discussion. This paper will examine the relationship of the psychological data to the educational practice of moral discussion.

Kohlberg (1978) has described what he terms the “psychologist’s fallacy”. He is referring to the fallacious assumption that what the psychologist acknowledges as empirical and theoretical “truth” can be assumed to have equal worth for the educator. Indeed, Kohlberg notes that he had to learn this the hard way; i.e., by making that fallacious assumption and suffering the consequence of an apathetic reaction by practitioners to his psychological prescriptions. In one way, this paper is open to the same criticism. We will be examining what psychology should actually prescribe *based on the psychological literature*. However, we will not be assuming that these prescriptions should be accepted unequivocally by the practitioner. That should be resolved through an ongoing dialogue between researchers and practitioners. Our goal here instead is merely to clarify what psychology actually has to offer. This is necessary because the literature has become quite murky concerning what the psychological research to date has to tell us about moral education. Examining and clarifying this is the first step in overcoming the psychologist’s fallacy. Only when an accurate perception of the data is available can we begin the dialogue with practitioners in order to make the research-based prescriptions meaningful for the educator.

While this paper will examine that aspect of moral education that is common to all varieties of application, i.e., moral discussion, it will focus most closely on one controversial aspect of moral discussion, what we term the “+ 1 convention”. This refers to the belief that, for moral discussion to successfully lead to the development of higher stages of moral reasoning, a discussion group requires a leader who will present reasoning one stage above the reasoning of the group members. In order to do this we will divide the presentation into

*Department of Psychology, Marquette University

three major parts: (1) an historical review of the psychological data concerning moral discussion and the +1 convention; (2) an analysis of the controversies surrounding, and actual uses of, the moral discussion techniques based on the psychological research findings; and (3) conclusions and prescriptions based on the preceding discussion. Let us begin with a review of the history of the research behind the moral discussion and +1 literature.

The psychological history of the +1 convention

The empirical investigation of the presentation of +1 reasoning began in the late 1960's with the work of Eliot Turiel and James Rest, two of Kohlberg's colleagues. The first experimental test of Kohlberg's stage theory was done by Eliot Turiel (1966). While the study still holds an important place in the literature supporting Kohlberg's theory, it has actually attracted a variety of strong criticisms for at least the past ten years (Broughton, 1978; Kurtines & Greif, 1974; Turiel, 1972; Zinn, 1970). Since such extensive treatments of the Turiel study are available elsewhere, only cursory review will be presented here. Turiel attempted to show that social modeling effects would fit the pattern predicted by stage theory, i.e., that only reasoning one stage above one's own would lead to change and that all change would be to the next highest stage. This is basically what Turiel concludes. Nevertheless, there are a number of problems in deriving such conclusions from Turiel's data. First, he finds very little, if any, actual development. Instead he chooses to subtract control group changes (regressions!) from experimental changes (stability). The result is *relative gain* by the experimental group, but *no actual development*. Second, he also finds relative regression by the -1 group. While these criticisms are not as simple as we would make them appear, it is important to realize the controversial status of the Turiel study. It alone certainly could not serve as the foundation for theory validation or application.

Furthermore, it is important to note that Turiel employed a simple one trial intervention. It is clear from the theoretical literature that one should not expect such a manipulation to have a developmental effect. Moral stage development is a slow process which includes both the maturation of the individual and his/her experience in a dialectical integration. Turiel (1969) recognized this in his later theoretical writings. Nevertheless, this early study was the foundation for much theory, research and application.

James Rest was interested in testing the theoretical assumptions of the theory explored by Turiel. He was interested in examining why only +1 reasoning was effective in causing development. Rest engaged in two studies of particular interest to this discussion. The first study was in collaboration with Turiel and Kohlberg (1969) and was in direct response to Turiel's 1966 study. The second was an elaborated follow-up of the first (Rest, 1973). While the findings of these two studies are actually quite complex, they have been represented quite simply and consistently in the literature.

Rest found that the subjects tended to rank low all responses at stages below their own, to restate responses at two or more stages above their own in terms of their own stage, and to prefer the response at one stage above their own to the response at their own stage. Thus Rest, like Turiel, was able to show that reasoning one stage above the subject's own level is what has most appeal for him (Hersh, Paolitto & Reimer, 1979, p. 104).

With regard to the hierarchical integration criterion, it has been demonstrated that adolescents exposed to written statements at each of the six stages comprehend or correctly put in their own words all statements at or below their own stage but fail to comprehend any statements more than one stage above their own. Some individuals comprehend the next stage above their own; some do not. Adolescents prefer (or rank as best) the highest stage they can comprehend (Kohlberg, 1978, p. 38.)

Most individuals can comprehend reasoning about moral issues at the next higher stage of reasoning above their present stage, but at no stage higher than that.... Individuals prefer reasoning at the highest stage they can comprehend (Beyer, 1978, p. 62).

These quotes are taken from the two most popular recent works on Kohlberg's moral

education.

The conclusions are then used to prescribe moral education curricula. Indeed, the last quote comes from an article entitled "Conducting moral discussions in the classroom". It is even more interesting to note that two of the most skeptical present writers about these empirical data are Rest (1980) and Turiel (1973) themselves. Rest has stated that "there is no evidence that '+1 modeling' is the effective condition for growth" (1980, p. 125) and suggested that the way comprehension was operationalized in his research would not be the way he would do it if replicating the study today (personal communication). Turiel (1973) has criticized his own study and has agreed with others (e.g. Broughton, 1978) in reinterpreting the original data.

It is most important to recognize that the conclusions quoted above are not accurate descriptions of the findings from the two Rest studies. While the literature tends to reflect Rest's findings as indicating that individuals prefer the highest stage they can comprehend, and comprehend one stage above their own stage, what Rest's studies actually demonstrate is that individuals *prefer the highest reasoning they are presented and comprehend the highest reasoning they already substantially produce*. Furthermore, reasoning higher than what they produce is either assimilated to their own stage, to +1 or is accurately represented. These actual findings deserve further elaboration.

First, preference is not for the +1 stage. It is typically for the highest stage presented, which for Rest's 1973 study was stage six. Unfortunately there is a further complication of this finding. Moran & Joniak (1979) found that the preference for higher stages of reasoning was a function of language complexity. When lower stage responses were worded more elaborately, they were preferred over higher stage statements. This finding severely questions the validity of even Rest's actual preference findings as well as the more popular interpretation of them.

Second, highest stage comprehension is not typically for the +1 stage as Beyer (1978) asserts above. Indeed Rest (1973) found

almost perfect comprehension of that stage (+1) by half of the subjects (called '+1 comprehenders') and the failure of the remaining half (the '0' or own-stage 'comprehenders') to comprehend +1 reasoning at all.... The highest stage used on the pretest, rather than the predominant stage, was the best predictor of comprehension.... All of the subject's responses are not scored at only one stage, however, and his use of stages higher than his own stage is of particular interest. Almost all of the subjects (84 percent) who showed on the pretest a substantial amount of codes at stages above their predominant stage also showed comprehension of statement above their predominant stage, whereas if a subject did not show substantial use of stages higher than his predominant stage there was only 30 percent probability of his comprehending statements higher than his predominant stage (p. 101).

Three additional research studies directly related to the question of the +1 manipulation have since been published. Tracy and Cross (1973) used an adaptation of Turiel's (1966) method of social modeling of +1 arguments on both sides of an issue. They found that exposure to +1 reasoning leads to development. They also found that this manipulation worked more effectively with preconventional subjects than with conventional subjects. It should be noted that Lockwood (1978) in his review of educational applications of moral development theory has raised the same issue of the relative appropriateness of different techniques for different stage individuals. One problem with the Tracy and Cross study is that they did not investigate exposure to other relative levels of reasoning (e.g., -1 or +2), so no comparative conclusions can be reached.

Keasey (1973) also investigated the effects of exposure to +1 reasoning. Keasey found that subjects who were posttested immediately (one day) demonstrated significantly more development if exposed to +1 reasoning than other conditions, but subjects given a delayed (two weeks) posttest showed no greater gains than other conditions. This may be due to the

weak interventions that were the norm during this phase of the literature. Like Turiel, these authors used one trial interventions, which, from a structural point of view, would be unlikely to have any effect.

Two other problems with these studies led to the next study we will report. First, these studies stage-typed subjects by their modal stage, ignoring other stages of reasoning the subjects might use. Second, all of these studies used prepared stereotyped moral arguments. Berkowitz, Gibbs and Broughton (in press) studied undergraduate moral dialogues and paired subjects on the basis of their modal and minor (20% or more of their total reasoning) stages. This was possible due to recent advances in the precision of the Kohlberg scoring system. Subjects were simply asked to discuss moral dilemmas presented on paper by the experimenter who was absent during the discussions. There was a series of five discussions over a span of approximately two months. The greater scoring precision allowed the investigators to demonstrate that the greatest development was for subjects paired with a partner one-third of a stage above them as opposed to two-thirds to a full stage, the same stage, or a non-discussion control group. One-third of a stage represents a pairing such as a pure Stage 3 subject with a subject predominantly at Stage 3 but with at least 20% Stage 4 or Stage 2 reasoning. The authors concluded that the optimal disparity of one-third was a refinement of the +1 convention and represented an optimal combination of shared and novel reasoning.

It is important to acknowledge at this point that there is a body of educational literature that tests the actual application of these findings to classroom practice (e.g., Blatt & Kohlberg, 1975; Colby et al., 1977). Lockwood reviews these studies and points out that

it should not be assumed, however, that plus-one matching is a well defined treatment; in practice it is extremely difficult to do. In actual practice, teachers probably express reasoning that is somewhat more sophisticated than that expressed by students or raise considerations relevant to the discussion that students have not taken into account. The point is that, for the studies employing plus-one matching, there is probably considerable variation in how the treatment is used (1978, p. 346).

Therefore we will not discuss these studies at this point. First we will examine how the educational literature actually prescribes classroom moral discussion. Included in this discussion will be the treatment of a number of controversial issues surrounding such implementation. Then we will return to some of the education studies in order to determine exactly what form the moral discussion took in those projects.

Review and critique of the educational literature on moral discussion

Now that we have examined the psychological basis for moral discussion, it is necessary to turn to the literature on implementation of these empirically derived principles in actual educational practice. In doing so we will be looking at three sources of information: (1) what the general theorists prescribe, (2) what the training manuals prescribe, and (3) what the practitioners actually do. In considering these sources, we will raise three points of controversy concerning moral discussion: (1) the role of teacher scoring skills, (2) the significance of the teacher's level of moral reasoning, and (3) the importance of class composition. First let us examine in general what the literature derives and prescribes from the psychological research discussed above.

The direct application of the above-cited psychological research is perhaps best exemplified by Downey and Kelly's statement that

when children are presented with arguments one stage above the level they have reached, the conflict or mismatch in the arguments will prompt them to attempt a resolution of the problem and thus help them towards the next stage of moral argument. It is this principle that both Kohlberg and Turiel hope will form a fundamental guideline for any programme of moral education in the school (1978, p. 81).

These authors explicitly base these conclusions on the Rest et al. (1969) and Turiel (1966)

studies. This position is quite common in the literature (cf. Beyer, 1978; Fenton, 1975; Fraenkel, 1977; Galbraith & Jones, 1976; Hersh & Mutterer, 1975; Selman, 1975).

Nevertheless, there is a great degree of diversity in the exact manifestations of this principle in different writers' treatments. Hersh & Mutterer (1975) argue that the teacher must know the child's present stage. Hersh, Paolitto & Reimer suggest that, while not necessary, scoring a child's stage may help one "become better listeners and respondents in discussions" (1979, p. 82). Galbraith and Jones (1976) on the other hand, argue that scoring is unnecessary. Downey and Kelly argue that "the main task then lies in increasing the ... level of stage mixture" (1978, p. 83). Some authors focus on peer interaction (e.g., Belenky, Tarule & Landa, 1979; Fraenkel, 1977) while others focus more on teacher facilitation (e.g., Fenton, 1975; Hersh & Mutterer, 1975). Regardless of the controversies involved, all authors agree that cognitive conflict engendered by classroom moral discussion leads to development via the exposure to reasoning one stage above one's own.

Thus teachers are instructed either to argue one stage above the majority of the class, paraphrase and highlight student +1 statements, insure a heterogeneous class mixture, or in other ways insure exposure to +1 arguments. Whatever the mechanism, the cognitive conflict is expected to lead to development to the next highest stage of reasoning, the goal of all such moral education programs.

In summary, there is both agreement concerning the goal of moral discussion and the general principles underlying it as well as disagreement concerning the exact form such discussion should take. Much of this confusion is due to the questionable empirical base for the prescriptions as well as common misreadings of those findings. At this point we will turn our attention to the three controversies cited above, i.e., scoring skills, teacher reasoning, and class composition. In considering each of these we will explore the general prescriptions just cited in more depth and will introduce some new information in order to attempt some resolution of those issues.

It should be noted first, however, that Kohlberg and many of his colleagues believe that classroom discussion is a useful but limited means of doing moral education. They (Hersh, Paolitto & Reimer, 1979; Kohlberg, 1978) argue that a "just community" approach to education is the necessary context for optimal student moral development (cf. Power, 1979). As noted above, moral discussion is at the heart of such more ambitious approaches also.

First let us examine what writers prescribe concerning the need for teachers to be able to score moral stages. Hersh and Mutterer (1975) argue that teachers must be able to identify the stage that each student is at in order to be able to facilitate classroom moral education. Hersh, Paolitto and Reimer (1979) spend thirty pages teaching the stages, but take a less absolute stance on the necessity of teacher scoring competency. They argue, however, that it helps a teacher as a facilitator to know the stages and the theory. Galbraith and Jones (1976) quite clearly argue against the need for a teacher to be able to score a child's stage of reasoning. In response to such suggestions, they respond,

No! The teaching process presented in this Handbook does not require teachers to become experts in identifying the stages of reasoning of their students... Teachers should not attempt to categorize individual students at particular stages. First of all, a single comment will not reveal an individual's modal stage of reasoning.... Second, in any classroom there will be a variety of modal stages of moral reasoning. Therefore, the teaching strategy outlined in the Handbook emphasizes small group and large group discussions where students at adjacent stages of development confront one another and seriously discuss genuine social and moral problems.... This is why teachers' skills in communication and group facilitation are more important to a productive discussion of a moral problem than their skill in stage interpretation (p. 184-5).

Fraenkel (1978) reacts to Rest's (1974) call for curriculum matching and the general call for +1 teacher strategies and argues that both are impractical because teachers will not

typically be one stage above all of their students, and even if they were it would be unlikely that they would have the time and expertise to process each student's reasoning, score it, and respond creatively one stage above it.

We therefore find two issues imbedded in this controversy. The first is whether teacher scoring skills are necessary, and the second is whether they are feasible. In explaining why the first half of their book for teachers is concerned with theory, Hersh, Paolitto and Reimer (1979) state that "Our experience convinces us that a careful approach to theory helps teachers generate more effective and more creative learning experiences for their students." They go on to explain that "We are skeptical of a 'bandwagon' approach to moral education. In leading in-service teacher-training workshops, we commonly find a premature rush to practice" (pp. 113-114). This author agrees with Hersh et al.'s middle ground stance and teaches in-service workshops with much the same philosophy in mind. The better a teacher understands the theoretical underpinnings of an educational program, the better s/he is able to adapt, problem-solve and innovate. S/he is a more flexible and autonomous implementer and therefore a more effective moral educator. Nevertheless, scoring moral stages is a difficult and complex task. Having scored literally more than a thousand moral interviews, having administered nearly as many, and having taught these skills to a wide variety of students and teachers, I recognize the impracticality of expecting classroom teachers to have readily mastered such skills. I have, however, found success in training some rudimentary, but useful, understanding of the stages and assessment techniques. As Galbraith and Jones suggest,

While teachers may associate general stage related remarks with particular stage reasoning, teachers should not attempt to categorize individual students at particular stages (1976, p. 185).

Hersh, Paolitto and Reimer (1979) concur:

It would be contrary to our teaching experience to expect that on first exposure a person could assimilate and accommodate all the information contained in the description of the stages.

But they do offer suggestions for improving mastery of such concepts with the

hope....that once readers learn to identify stage-related arguments, they can become better listeners and respondents in discussions involving moral issues. That, in turn, is an important step in becoming more effective moral educators (p. 82).

In other words, teacher training is imperative but teacher mastery of stage scoring is a useful luxury. In their approaches to moral education, Belenky, Tarule and Landa (1979), and Hersh and Mutterer (1975) highlight teacher training. However, it is important to focus on what the essential aspects of teacher training are. Fraenkel's (1977) critique is predicated on the assumption that +1 inputs need to come from the teacher alone. Belenky et al. (1979) and Galbraith and Jones (1976) make it clear, however, that students must share in those inputs. Thus what Fraenkel terms impractical may also be viewed as unnecessary. This brings us to the next issue.

A second controversial question is whether teachers need to be post-conventional moral reasoners in order to be effective moral educators? While no one has claimed that teachers need to be members of a moral elite, one may infer from portions of the prior discussion that it would be desirable or even necessary if teachers are to learn and use the theory and stages. Indeed, Jack Fraenkel (1976, 1977, 1978), in his oft reprinted critique of Kohlberg psychology and education, has concluded that implicit in the model is the mandate for teachers to be more developed than *all* of their students. Fraenkel has argued that

since Kohlberg has stated that only ten percent of the population reaches Stages 5 or 6, the laws of probability suggest that there are many teachers who themselves reason at the lower stages, and who accordingly are likely to come in contact with students reasoning at stages higher than their own (1978, p. 254).

I would question both the logic in Fraenkel's contention as well as his empirical basis.

Kohlberg has not only argued that most adults (80 to 90%) do not develop beyond the conventional level of moral reasoning, but also that

empirical research between 1968 and 1976 did not confirm my theoretical statements about a sixth and highest stage. My longitudinal subjects, still adolescents in 1968, had come to adulthood by 1976, but none had reached the sixth stage. Perhaps all the sixth stage persons of the 1960s had been wiped out, perhaps they regressed, or maybe it was all my imagination in the first place (1980, p. 457).

Earlier Kohlberg had argued that

fully-principled or Stage 5 and especially Stage 6 thinking is an adult development, typically not reached until the late twenties or later (1973, p. 190).

We thus can not only conclude that teachers are not likely to be post-conventional reasoners, but also that students are definitely reasoning no higher than Stage 4. Furthermore, since 95% of adults are Stage 3 or higher (Candee, Graham & Kohlberg, 1978), it is unlikely that many teachers will encounter students reasoning at a stage higher than themselves. From my own research it has become clear that the vast majority of college students reason no higher than Stage 4; e.g., only 1 out of 82 undergraduates in one study (Berkowitz, Gibbs & Broughton, in press) was higher than Stage 4 and that individual was lower on a posttest followup. Of these same subjects, only 17 had more than a beginning minor amount of Stage 4 reasoning and only three had fully abandoned Stage 3 reasoning. In one extensive study of high school student reasoning (Power, 1979), only 22 of 104 students in a *just community school* ever achieved any Stage 4 reasoning. In another massive study, only 4 out of 235 junior and senior high school students had developed any Stage 4 reasoning at all, and in a sample of an additional 123 eleventh grade "slow learners" in the same study, there was no evidence of any Stage 4 reasoning (Colby, Kohlberg, Fenton, Speicher-Dubin & Lieberman, 1977).

It is thus unlikely that many high school teachers will encounter students reasoning higher than the teacher. Furthermore, it is not accurate to argue that a teacher must be at the student's stage to comprehend and reproduce it. We know that the teacher merely needs to be able to produce some reasoning at the next highest stage in order to be able to comprehend and reproduce it (Rest, 1973). Thus, if a student is at Stage 4, the teacher need merely be at Stage 3(4) to have the competency to do Kohlberg moral discussion with the student. Most students will not be at Stage 4 and few teachers will not have reached Stage 3(4).

A further point that should be made along these same lines is that development seems to continue as long as education continues (Rest, 1979). Teachers have at least four more years of education and, consequently, moral development, than their students. This would explain part of the relative absence of an overlap between the range of stage scores for teachers and students.

Wilkins (1980) has tried to study the actual degree of overlap between the two groups. Since I have reviewed this research elsewhere (Berkowitz, in press), I will discuss it only briefly here. Wilkins concluded that teacher reasoning overlaps student reasoning enough that many teachers will encounter students who reason at a stage higher than their own. There are two basic problems with Wilkins' study: the measure of moral reasoning chosen and the samples studied. Wilkins used Rest's (1979) Defining Issues Test (DIT) as his measure of moral reasoning. The DIT is a measure of one's capacity to *evaluate presented reasoning*. Wilkins and Fraenkel are conceptually interested in the ability of teachers to *spontaneously produce* higher stage *reasoning*, not merely recognize and evaluate it once presented. Kohlberg's Standard Moral Judgment Interview measures production capacity. It therefore seems more appropriate to use the Kohlberg measure to assess the relative levels of development of teachers and students for this empirical question. The above-cited data from a variety of studies serves to suggest that the degree of overlap found by Wilkins may

be misleading.

The samples studied by Wilkins further confound his results. He used a sample of American high school students studied initially by Rest and a sample of education students in Australia. He then argued that the overlap in distributions of reasoning between the two groups indicates that actual classroom teachers will overlap in levels of development with their students. Given Rest's (1979) clear finding that development continues as long as education continues, it seems somewhat misleading to study education students rather than actual teachers. In a study of American educators, Griffiore and Lewis (1978) report slightly higher DIT scores than Wilkins' education students. Furthermore, Wilkins used Rest's P score which is a measure of Stages 5 and 6. We already know that the overlap of teachers and students will be at Stages 3 and 4, not 5 and 6. The DIT does inflate one's stage score by one or two stages (Griffiore and Lewis found teachers to be predominantly at Stages 4 and 5), but conceptually we are interested in what stages the students can produce that their teachers may not be able to. In conclusion, there seems to be no logical or empirical rationale in support of Fraenkel's critique.

The third and last controversy concerns the requirements for class composition. In other words, does the composition of the class relate in any way to the success of a moral discussion program. Certainly, there are many variables that might seem relevant here, such as verbal skills, cognitive development, etc. We will focus on the most direct and relevant variable, i.e., students' stages of moral reasoning. One often overlooked aspect of classroom moral discussion is the degree of variance of reasoning represented by the levels of moral development of the students. One of the earliest indications that this might be relevant was the findings of the Colby et al. (1977) study. These authors found that the success (in terms of average moral development) of their classroom interventions was largely determined by the degree of mixture of students' stages of moral reasoning. Heterogeneous classes were more likely to evidence development than were homogeneous classes.

In a more recent study of classroom moral discussion curricula, the same finding has been reported as quite pronounced (Gibbs, personal communication). Berkowitz, Gibbs and Broughton (in press), as noted above, found that heterogeneous dyads engendered more development than same-stage dyads. In the parallel study of dyads engaged in logical, non-moral tasks, Perret-Clermont (1980) concludes that heterogeneous dyads seem to promote more development than homogeneous dyads. It thus seems that classroom heterogeneity in terms of moral stages is an important consideration for a moral educator. The full significance of this conclusion will become evident when we finally try to integrate all of the points made thus far. First however, now that we have considered what is prescribed for moral educators, we will turn to an examination of what moral educators actually do in moral discussions.

How is moral discussion actually implemented? When we look at the two major texts that can be considered teacher training manuals (Galbraith & Jones, 1976; Hersh, Paolitto & Reimer, 1979), we find that moral discussion is more a matter of group facilitation than the +1 modeling originally envisioned by students of the work of Turiel and Rest. Galbraith and Jones note that

a teacher who facilitates a discussion of a moral dilemma has two primary tasks: promoting student interaction and making certain that the discussion remains focused on the moral issues of the story (1976, p. 46).

These authors therefore train a variety of probe questions as required teacher behavior. Hersh, Paolitto and Reimer similarly describe in detail the skills teachers need in order to effectively lead moral discussions. These skills include creating cognitive conflict, stimulating student perspective-taking, and using a variety of questioning strategies. Thus, while

they advocate learning the stages of moral reasoning, and while they go to great lengths to teach the stages, they do not try to train +1 strategies.

Lockwood (1978), as noted above, cautions that +1 matching is difficult to achieve and should not be assumed to have occurred in a moral discussion program. Fraenkel (1978) claims that it is not even feasible in many instances. Kohlberg (personal communication) has suggested that we analyze the transcripts of moral discussions printed in Blatt and Kohlberg (1975). Doing so may offer us an insight into exactly what the teacher is actually doing in moral discussions.

Blatt and Kohlberg include, as an appendix to their study, excerpts of two moral discussions led by the first author. These cases are not cited as exemplary, rather merely as illustrative of Blatt's discussions. In inspecting these transcripts, one sees that Blatt is clearly trying to highlight important moral issues in the dilemmas under discussion, he is more apt to Socratically point out contradictions, offer exaggerated paraphrases or suggest problematic considerations than to present a clear higher stage argument. While some of Blatt's statements may be construed as at a particular stage, none are elaborated sufficiently to be clearly scored. We can therefore see that +1 is not the primary feature of Blatt's discussion behavior.

Other authors have also printed excerpts of moral discussions. Hersh, Paolitto and Reimer (1979) present a segment in which there are no attempts to produce +1 arguments. The teacher simply acts as a facilitator and elicitor. Scharf (1978) offers a similar transcript of moral discussion by a group of delinquent youth. Here also the leader is basically facilitating rather than offering +1 arguments. Finally, Colby et al. (1977) present two lengthy discussions as an appendix to their presentation. The leader is only identified in the first discussion. In this case, Ted Fenton is leading eighth grade students as a demonstration for a teacher. While many student statements are scored for stage by the authors, none of Fenton's statements are scored or scorable. Fenton is basically probing and orchestrating the group discussion. It therefore seems that +1 techniques are neither trained nor used in actual moral discussion programs.

Conclusions and Revisions

We have raised a number of issues related to moral discussion and, more specifically, the +1 convention. We have examined these issues and the relevant data at some length. Now we must look back over this discussion and attempt to reach some conclusions based on our analysis. These conclusions will serve to revise and consolidate the +1 convention by bringing it into alignment with the actual data in the literature as well as by achieving some consensus from the diverse interpretations that we have shown to flourish in the literature. Because we have raised so many controversies and questioned the validity of much data, this will not be a simple task.

It is probably best to begin by pointing out that the criticisms of the psychological data typically offered as "proof" of the +1 phenomenon should not lead one to thoroughly reject the validity of those studies. While they each have their problems, when taken together we may be able to uncover a pattern of results that informs our understanding of the psychological processes underlying moral discussion. It is clear that the most common finding is that exposure to higher stage reasoning, roughly within one stage of one's own stage, is related to the subsequent development of the individual. Some findings argue for a strict +1 conclusion. Others argue for variations of that. We may conclude, however, that within a given range of stage disparities, it is recommended that students are exposed to higher stage reasoning.

A second issue concerns the source of those higher stage arguments. Three sources are

observed in the literature: teachers' verbal behavior, peer verbal behavior, or written argument. No one has studied the long term impact of written moral arguments, so we have no basis for assuming that such a format would be successful. We have, however, reviewed much evidence concerning the relative roles of student and teacher inputs. It should be clear that peer reasoning seems to be more influential in promoting moral development through group discussion than is teacher reasoning. Colby et al. (1977) found that classroom heterogeneity was a central factor in determining the success of a moral discussion program in conjunction with teacher facilitation behavior, not teacher stage of reasoning. Blatt and Kohlberg (1975) found that middle class children benefited almost as much from leaderless discussions as from discussions led by a teacher. Berkowitz et al. (in press) found undergraduate dyadic discussion to be roughly as effective as studies with guided discussion procedures. Finally, when we reviewed what the training manuals prescribe and what the classroom practitioners actually do, we discover that teacher + 1 inputs were virtually non-existent. It is therefore concluded that students should be exposed to the higher stage arguments of their peers in moral discussion.

What then is the role of the classroom teacher? As we have already noted, beyond the controversy regarding the +1 convention, there seems to be general agreement in the literature that teachers should facilitate the class interaction in moral discussion. This is still somewhat vague however. What form should this facilitation take? In looking at the two published "training manuals" (Galbraith & Jones, 1976; Hersh, Paolitto & Reimer, 1979), we noted that a fundamental task assigned to the teacher was questioning. Indeed, Colby et al. (1977) also suggested that a Socratic mode of teaching was a primary characteristic of the successful practitioner of classroom moral discussion. We may then ask why such a teacher mode is effective. We know that most authors agree that cognitive conflict, engendered through moral discussion, will lead to individual moral development. This is the position of both the theorists and the researchers. It makes sense that teacher questioning will produce cognitive conflict in the students by raising contradictions or other problematic concerns. Gibbs (1980) has reviewed the literature on teacher training for classroom moral discussion and has culled from these sources (predominantly Galbraith & Jones, 1976; Gomberg, Cameron, Fenton, Furtek & Hill, 1980) what he considers to be the best guidelines for discussion leadership. He notes that the teacher needs to assume the responsibility for maintaining an exchange of ideas that are focused on the problem at hand. Furthermore, they should refrain, when possible, from offering their point of view. Finally, Gibbs points out that teacher questioning should not be reduced to "interviewing" of individual students. Instead teacher facilitation should lead to student interaction. It is not within the scope of this paper to detail the variety of specific teacher behaviors that are advocated for the facilitation of classroom moral discussion. For this we refer the reader to the sources cited in the above discussions. It is important, however, to note some of the general categories of teacher behavior that are advocated. Questioning is certainly important. Questions may serve an elicitive function (i.e., eliciting further information or clarification) or a challenging function (e.g., pointing out contradictions or inconsistencies). Teachers may also paraphrase student reasoning. It is also important for the teacher to engage in task-oriented statements, such as keeping a moral focus or clarifying the dilemma under discussion. Finally, it is important for the teacher to engage in behavior which affects the climate of the classroom and subsequently the moral discussion. Such behavior may include statements that promote acceptance of others' reasoning or an orientation to fairness in interpersonal interactions. While this is not meant to be an exhaustive categorization, it does represent the major issues in moral discussion leadership.

There is one further ingredient that we would care to mention. It is one that is often overlooked in the rush to train teachers to be effective moral discussion leaders, but one which most practitioners will probably endorse. In her interview in *Promoting Moral*

Growth, (Hersh, Paolitto & Reimer, 1979), Margot Strom includes as one of five fundamental guidelines for discussion teachers the need for "developing student skills". All too often it is assumed that an effective teacher is all that is needed to produce an effective classroom discussion. Strom is wise in her inclusion of student skills. If, as we argue above, it is the peer interaction that is most significant in leading to individual development, and, if the teacher is primarily supposed to be facilitating this interaction, it stands to reason that student communication and role taking skills are vital to the success of such an enterprise. Much of the research done by Ralph Mosher, Norman Sprinthall and their students (e.g., Sprinthall, 1980; Sprinthall & Mosher, 1978; cf. Lockwood, 1978) has focused on training listening and communication skills in students and relating those skills to the moral development of their subjects.

This author has spent many years studying the nature of moral discussion and has derived a set of eighteen "transactive" behaviors that relate significantly to the success of a discussion in leading to moral development. (Berkowitz, 1980a, 1980b; Berkowitz & Gibbs, 1979; Berkowitz, Gibbs & Broughton, 1980). It is our contention that all people do not naturally engage in all of those skills and that some people engage in very few. If we could train such behavior in students, it should lead to richer, more interpenetrative and successful moral discussions. The current teacher training models therefore should train teachers to be able to train their students in the appropriate discussion skills.

In summary then we can make the following prescriptions for classroom moral discussion. First, class composition is vital. A teacher will experience more success if there is a degree of heterogeneity of student moral reasoning levels. Student skills are also vital. It is useful to train student communication skills. Certainly, teacher behavior is important. The teacher should create a suitable atmosphere for the open discussion of moral issues and should facilitate the student interaction that is the core of that discussion. We do not believe that teachers need to argue one stage above the stage of the reasoning of students in the class. Peers in a heterogeneous class will provide such reasoning. It is encouraging to note that, although such has not been demonstrated in the moral development literature, Perret-Clermont (1980) has found that higher stage peers also develop as a product of peer discussion of logical tasks.

These guidelines are certainly neither specific nor perfect enough to serve as a recipe for moral discussion. That was not our intention. Another reason for this is that there are many studies that yet need to be done to clarify some of the vagueness in those prescriptions. We will therefore close this paper with some suggestions for future research; research that needs to be done in order for moral discussion to leave its adolescence and reach maturity.

One implication of this review for research is to encourage further investigations of the nature of moral discussion, particularly its transactive features. Another implication is that we need to investigate the exact parameters of class composition that relate to successful intervention programs. It is clear that heterogeneity of reasoning is necessary, but beyond this we have no specific prescriptions. Lockwood (1978) suggests we explore also the differential effects of moral discussion on students at different stages. Blatt and Kohlberg (1975) have indirectly raised the question of sex as an intervening variable. Regardless of which of these issues we focus on, the task before us is the further refinement of the goals and means of moral discussion. It is our hope that this paper will promote the discussion of those issues we have raised.

Résumé

La discussion morale est au centre de l'enseignement de la morale, dont une partie importante provient de l'école cognitive de psychologie.

Une grande partie des ouvrages sur l'enseignement de la morale est basée sur la recherche psychologique émanant du modèle théorique. Dans cet article, nous examinerons le rapport entre cette recherche et la discussion morale. C'est par une revue critique de la recherche et par les recommandations pédagogiques qui en dérivent que nous illustrerons un aspect particulièrement contesté de la discussion morale, à savoir la convention "+ 1". Nous baserons nos conclusions sur cette analyse.

Footnotes

- Belenky, M. F., Tarule, J. M. and Landa, A. M. *Education and development: The role of schools in ego-identity and moral development of normal and troubled adolescents*. Plainfield, Vermont: The Goddard/Teacher Corp Institute in Developmental Education, 1979.
- Berkowitz, M. W. The role of transactive discussion in moral development: The history of a six-year program of research-Part I. *Moral Education Forum*, 1980a, 5, 13-26.
- Berkowitz, M. W. The role of transactive discussion in moral development: Description and applications of the transactive model-Part II. *Moral Education Forum*, 1980b, 5, 15-27.
- Berkowitz, M. W. Moral peers to the rescue! A critical appraisal of the "+ 1" convention in moral education. *Phi Delta Kappan*, in press.
- Berkowitz, M. W. and Gibbs, J. D. *A preliminary manual for coding transactive features of dyadic discussion*. Unpublished manuscript, Marquette University, Milwaukee, Wisconsin, 1979.
- Berkowitz, M. W., Gibbs, J. C., and Broughton, J. M. Structure and process: An experimental inquiry into the "what" and "how" of moral judgment development. In *Proceedings of the Eight Annual International Conference on Piagetian Theory and its Implications for the Helping Professions* (Vol. 1). Los Angeles: University of Southern California Press, 1980.
- Berkowitz, M. W., Gibbs, J. C. and Broughton, J. M. The relation of moral judgment stage disparity to developmental effects of peer dialogues. *Merrill-Palmer Quarterly*, in press.
- Beyer, Barry K. conducting moral discussions in the classroom. In P. Scharf (Ed.), *Readings in Moral Education*. Minneapolis: Winston Press, Inc., 1978.
- Blatt, M. and Kohlberg, L. The effects of classroom moral discussion upon children's level of moral judgment. *Journal of Moral Education*, 1975, 4, 2, 129-161.
- Broughton, M. M. The cognitive-developmental approach to morality: a reply to Kurtines and Greif. *Journal of Moral Education*, 1978, 7, 81-96.
- Candee, D., Graham, R., and Kohlberg, L. *Moral development and life outcomes* (NIE-6-74-0096). Cambridge, Massachusetts: Harvard University, Center for Moral Development and Education, 1978.
- Colby, A., Kohlberg, L., Fenton, E., Speicher-Dubin, B., and Lieberman, M. Secondary school moral discussion programmes led by social studies teachers. *Journal of Moral Education*, 1977, 6, p. 90-111.
- Downey, M. and Kelly, A. V. *Moral Education: Theory and Practice*. NY: Harper and Row, 1978.
- Fenton, E. A developmental approach to civic education. In J. Meyer, B. Burnham and J. Cholvat (Eds.), *Values education: Theory/practice/problems/prospects*. Waterloo, Ontario, Canada: Wilfrid Laurier University Press, 1975.
- Fraenkel, J. R. The Kohlberg bandwagon: Some reservations. *Social Education*, April 1976, 216-222.
- Fraenkel, J. R. *How to teach about values*. Englewood cliffs, NJ: Prentice-Hall, Inc., 1977.
- Fraenkel, J. R. The Kohlberg bandwagon: some reservations. In P. Scharf (Ed.), *Readings in Moral Education* Minneapolis: Winston Press, 1978.
- Galbraith, R. E. and Jones, T. M. *Moral reasoning: A teaching handbook for adapting Kohlberg to the classroom*. Minneapolis: Greenhaven Press, Inc., 1976.
- Gibbs, J. C. *Dilemma discussion instructional format*. Unpublished paper, Ohio State University, Columbus, Ohio, 1980.
- Gomberg, S. H., Cameron, D. K., Fenton, E., Furtek, J. and Hill, C. L. *Leading dilemma discussions: A workshop*. Unpublished manuscript, Carnegie-Mellon University, Pittsburgh, PA, 1980.
- Griffore, R. J. and Lewis, J. characteristics of teachers' moral judgment. *Education Research Quarterly*, Fall 1978, 3, 20-30.

- Hennessy, T. C. (Ed.) *Values and Moral Development*. N.Y.: Paulist Press, 1976.
- Hennessy, T. C. *Value/moral education: Schools and teachers*. N.Y.: Paulist Press, 1979.
- I'ersh, R. and Mutterer, M. Moral education and the need for teacher preparation. In J. Meyer, B. Burnham and J. Cholvat (Eds.), *Values education: Theory/practice/problems/prospects*. Waterloo, Ontario, Canada: Wilfred Laurier University Press, 1975.
- Hersh, R., Paolitto, D. and Reimer, J. *Promoting moral growth: From Piaget to Kohlberg*. New York: Longman, 1979.
- Keasey, C. B. Experimentally induced changes in moral opinions & reasoning. *Journal of Personality & Social Psychology*, 1973, 26, 30-38.
- Kohlberg, L. Continuities in childhood and adult moral development revisited. In P. B. Baltes and K. W. Schaie (Eds.), *Life-span developmental psychology: Personality & socialization*. N.Y.: Academic Press, 1973.
- Kohlberg, L. Moral stages and moralization: The cognitive-developmental approach. In T. Lickona (Ed.), *Moral development and behavior: Theory, research, and social issues*. New York: Holt, Rinehart, & Winston, 1976.
- Kohlberg, L. The cognitive-developmental approach to moral education. In P. Scharf (Ed.), *Reading in Moral Education*. Minneapolis: Winston Press, 1978.
- Kohlberg, L. Educating for a just society: A updated and revised statement. In B. Munsey (Ed.), *Moral development, moral education, and Kohlberg: Basic issues in philosophy, psychology, religion, and education*. Birmingham, Alabama: Religious Education Press, 1980.
- Kohlberg, L. and Wasserman, E. R. The cognitive-development approach and the practicing counselor: An opportunity for counselors to rethink their roles. *The Personnel and Guidance Journal*, May 1980, 559-567.
- Kurtines, W. and Greif, E. B. The development of moral thought: Review and evaluation of Kohlberg's approach. *Psychological Bulletin*, 1974, 81, 453-470.
- Lockwood, A. L. The effects of values clarification and moral development curricula on school-age subjects: A critical review of recent research. *Review of Educational Research*, 1978, 48, 325-364.
- Maitland, K. A. and Goldman, J. R. Moral judgment as a function of peer group interaction. *Journal of Personality and Social Psychology*, 1974, 30, 699-704.
- Moran, J. J. and Joniak, A. J. Effect of language on preference for responses to a moral dilemma. *Developmental Psychology*, 1979, 15, 337-338.
- Perret-Clermont, A. *Social interaction and cognitive development in children*. New York: Academic Press, 1980.
- Power, C. *The moral atmosphere of a just community high school: A four year longitudinal study*. Unpublished doctoral dissertation, Harvard University, 1979.
- Power, C. and Reimer, J. Moral atmosphere: An educational bridge between moral judgment and action. In W. Damon (Ed.), *New Directions for Child Development: Moral Development*. San Francisco: Jossey Bass, in press.
- Rest, J. Patterns of preference and comprehension in moral judgment, *Journal of Personality*, 1973, 41, 86-109.
- Rest, J. Developmental psychology as a guide to value education: A review of "Kohlbergian" programs. *Review of Educational Research*, Spring 1974, 241-259.
- Rest, J. R. *Development in Judging Moral Issues*. Minneapolis: University of Minnesota Press, 1979.
- Rest, J. Developmental psychology and value education. In B. Munsey (Ed.), *Moral development, moral education, and Kohlberg: Basic issues in philosophy, psychology, religion, and education*. Birmingham, Alabama: Religious Education Press, 1980.
- Rest, J., Turiel, E. and Kohlberg, L. Relations between level of moral judgment and preference and comprehension of the moral judgment of others. *Journal of Personality*, 1969, 37, 225-252.
- Scharf, P. *Moral Education*. Davis, California: Responsible Action Press, 1978.
- Selman, R. A. Developmental approach to interpersonal and moral awareness in young children: Some theoretical and educational implications of levels of social perspective taking. In J. Meyer, B. Burnham and J. Cholvat (Eds.), *Values education: Theory/practice/problems/prospects*. Waterloo, Ontario, Canada: Wilfrid Laurier University-Press, 1975.
- Sprinthall, N. A. Psychology for secondary schools: The saber-tooth curriculum revisited? *American Psychologist*, 1980, 35, 336-347.

- Sprinthall, N. A. and Mosher, R. L. *Value development as the aim of education*. Schenectady, New York: Character Research Press, 1978.
- Tracy, J. J. and Cross, H. J. Antecedents of shift in moral judgment. *Journal of Personality and Social Psychology*, 1973, 26, 238-44.
- Turiel, E. An experimental test of the sequentiality of developmental stages in the child's moral judgments. *Journal of Personality and Social Psychology*, 1966, 3, 611-618.
- Turiel, E. Developmental processes in children's moral thinking. In P. Mussen, J. Langer and M. Covington (Ed.), *Trends and issues in developmental psychology*. New York: Holt, Rinehart, and Winston, 1969, 92-133.
- Turiel, E. Stage transition in moral development. In R. M. Travers (Ed.), *Second handbook of research on teaching*. Chicago: Rand McNally, 1972.
- Turiel, E. The effects of cognitive conflicts on moral judgment development. Unpublished manuscript, Harvard University, Cambridge, MA, 1973.
- Wasserman, E. R. Implementing Kohlberg's "Just Community Concept" in an alternative high school. In P. Scharf (Ed.), *Readings in moral education*. Minneapolis: Winston Press, 1978.
- Wilkins, R. A. If the moral reasoning of teachers is deficient, what hope for pupils? *Phi Delta Kappan*, April 1980.
- Zinn, V. Methodological issues in the study of the development of moral judgment. Unpublished manuscript, Clark University, Worcester, Massachusetts, 1970.