Teaching of Critical Thinking/Thinking—Promises! Promises!

JOHN FOLLMAN  University of South Florida

Critical Thinking Bandwagon

There has recently been a "new wave," an explosion, of interest, in the construct of critical thinking/thinking. Critical thinking/thinking activity had been in repose for years, until about 1983, but is now a bandwagon if not a juggernaut.

Sternberg (1985) addressed the force behind this "new wave" critical thinking bandwagon:

"Probably never before in the history of educational practice has there been a greater push to teach children to think critically."

Critical thinking has now been broadened to include thinking (Sternberg, 1985, a); informal reasoning (Perkins, 1985); scientific thinking (Tweney, 1987); framework of thinking (Rankin & Hughes, 1986); classroom discourse (Hultgren, 1987); classroom discussion (Hauser, 1987); even learning (Bransford & Stein, 1984) and intelligence (Baron, 1985).

Objectives

Three objectives are addressed in this paper. The main objective is to document the difficulty of teaching critical thinking/thinking. The second objective is to articulate the conditions necessary for success in teaching critical thinking/thinking. The third objective is to indicate the contemplated consequences if these conditions are not met.

Effects of Contemporary "New Wave" Critical Thinking/Thinking Skills Training Programs

In her overview of thinking instruction programs, including some of the "new wave" models, i.e., Covington's Productive Thinking Program, and de Bono's CORT program, Quellmalz (1983) concluded that "most of the programs reviewed have not succeeded." She also concluded that data documenting a critical thinking instruction program's implementation and effectiveness is often absent, incomplete, or based on poorly developed or selected measures.

Gagne (1983) overviewed the teaching of thinking skills literature and concluded that there is little to be learned from most of the research on teaching critical thinking skills.

Eight studies all indicating that the Lipman Philosophy for Children program was effective in enhancing student's thinking skills, were overviewed, mostly anecdotally by Johnson (1984).

Sternberg & Bhana (1986) recently evaluated the effects of "new wave," thinking skills training programs by synthesizing the "research" on five well-known, widely used, "new wave" thinking skills training programs: Instrumental Enrichment (Feuerstein, 1980)—38 "studies"; Philosophy for Children (Lipman, Sharp & Oscanyan, 1980)—20 "studies"; Structure of the Intellect (Meeker, 1969)—36 "studies"; Problem Solving and Comprehension (Whimbey & Lochhead, 1979)—three "studies"; and Odyssey (Herrnstein, Nickerson, Sanchez & Swets (1986) reported in Sternberg &
Bhana (1986)—one “study”), all referenced in Sternberg & Bhana. The major conclusion of this synthesis was that “some thinking skills training programs are probably not a whole lot better than snake oil, but the good ones, although not miracle cures, may improve thinking skills.” The equivocal, tentative, character of this conclusion is instructive. Sternberg & Bhana (1986) also concluded “although all but a few of the available evaluations leave a great deal to be desired, there are enough positive results to suggest the potential for gains.”

Bransford, Burns, Delclos & Vye (1986) concluded that Sternberg & Bhana’s (1986) analysis of the effects of published thinking skills programs is sobering but valid. They then cited some other, highly limited, highly selected, studies that demonstrated effective teaching of what they characterised as a problem solving approach to teaching concepts and skills. See also Bransford, Sherwood, Vye & Rieser (1986) for an elaborated version of this article in which they link teaching of reasoning, thinking and problem solving.

In a treatment study of the effects of education on informal logic on secondary, undergraduate and graduate students, Perkins (1986) found a slight but disappointing impact. A “scaffolding” pedagogical approach manifested some promise.

Savell, Twohig & Rachford (1986) reviewed the empirical research on the effects of Feuerstein’s Instrumental Enrichment (FIE) as a method of teaching thinking in four countries. After noting that many of the studies failed to find clear FIE effects and that most studies were difficult to interpret they also concluded that there was a subset of data suggesting that FIE might be having an effect although it was not clear what that effect meant.

### Overview of the Effectiveness of “New Wave” Thinking Skills Training Programs

For those educators who wish these critical thinking/thinking skill training programs to be effective the operative words are Promises! Promises! For those educators who require more evidence than Sternberg & Bhana’s (1986) “potential” there is always the old research from which to draw leads about the prospects for the “new wave” critical thinking skill training programs.

Since independent, methodologically sound, replicated, research remains to be conducted on the effects of the critical thinking/thinking skill instruction “new wave” programs, it may be tuitional to look at the effects of some of the pre “new wave” critical thinking instruction research.

#### Effects of Pre “New Wave” Critical Thinking Skill Training Programs

Reviewed are the major conclusions and implications of a pre “new wave” (early 1970’s) critical thinking research program conducted by the author and colleagues, as well as related research by other, independent, researchers, in order to address the prospects of the present “new wave” critical thinking skill training programs.

This critical thinking research program consisted of six different streams of critical thinking research, overview of which suggests the following conclusions:

1. Most of the recent reviews of the literature are, at best, superficial.
2. A number of psychometrically satisfactory instruments are available and/or are in development (Follman, in press).
3. The construct of critical thinking appears to have little, if any, unique variance (Follman, in progress).
4. Instruction of critical thinking is not easy.

In 1971 Skinner (1971) observed that teaching critical thinking is a myth.

Initially presented is an update of the critical thinking literature over the last,
ca. 10-15 years, immediately prior to the “new wave,” i.e., from ca. 1968 to ca. 1983.

The developing of critical thinking in students as an academic objective has proved to be very challenging. The most reasonable generalization from the old, ca. 1968 to ca. 1983, research cited herein re successful teaching of critical thinking is that while critical thinking has occasionally been developed, enhanced, improved, taught, learned, etc., it has not been either consistent or common. When critical thinking has successfully been developed, it has often been with younger, less sophisticated subjects, especially middle school age level or younger, rather than with older, more verbal, more sophisticated students. Many of the treatments have been unrealistically brief, too brief to be beneficial. It is noted that few studies were reported during this period in which critical thinking was successfully taught. Six pre-college age level studies did report some significant, although mixed treatment effects [Coble & Hounshell (1972); Curtis (1980); George & Dietz (1968); Hunkins (1970); Mathias (1973); Shipman, Jansson & Heimer (1974)]. Across 10 college age level studies, at best mixed results obtained [Annis & Annis (1979); Browne, Haas & Keeley (1978); Ferris (1973); Keeley, Browne & Kreutzer (1982); Logan (1976); Moll & Allen (1982); Ross & Semb (1981)—gains confounded with mastery approach; Smith (1977); Story (1974); Wells, Yopp, Mend & Bean (1982)].

The mostly prior to 1950 literature of methods of teaching critical thinking in secondary schools and also elementary schools are overviewed by Burton (1960). On the secondary level, positive treatment effects were found in 11 studies. On the elementary level, 11 articles were reported about what might be effective teaching methods but did not report testing, one study reported a testing program but was not soundly designed, and another study was on such a large scale that Burton was unable to determine details of effectiveness. Burton listed the methods of teaching critical thinking he considered established in the studies above, most of which were in biology, chemistry, geometry, algebra, and some in language arts. Most of these methods were instruction of aspects of scientific method and experimental design, almost generically.

Russell (1960) reported in the Encyclopedia of Educational Research that “the ‘teachability’ of critical thinking had been demonstrated in several curricular areas.” However, the date and the type of terminology used herein, i.e., “demonstrated,” suggest the extant lack of sophistication and of rigor of the research methodology.

Teaching of critical thinking in social studies was overviewed by Shaver (1962) who concluded that the primary basis for the choice of treatments had to be based on intuition, not on the criterial educational research.

Young (1966) reported a considerable number of examples of studies in which critical thinking was taught or strategies to teach it were successful. Most of these efforts would be seen today as pilot studies and/or demonstration projects.

In an overview of a number of efforts to teach critical thinking viewed as a process, Saadeh (1969) concluded that the evidence for the development of critical thinking was overwhelming.

Lewis & Dahl (1970) overviewed at least 15 “surveys” concluding that the efficiency of any one method of enhancing critical thinking was still largely conjectural. They also concluded that few studies found enhancement of critical thinking outside the scientific disciplines.

Shipman, Jansson & Heimer (1974) reported that there had been several studies, four were actually cited, that had attempted (“with little success”) to teach critical thinking.

Research on inquiry/discovery teaching in social studies over the January 1967-November 1972 period was examined by Marsh (1974). Of 28 studies,
11 reported significant treatment success. Marsh concluded that all of these 11 studies were methodologically unsound. He likewise criticized the other 17 as methodologically unsound.

Wright (1977) asserted that relatively little research had focused on methods of teaching critical thinking.

Wulf & Garett (1978) linked critical thinking to the formal operations period of Piaget's developmental framework and reported a limited review of effects of critical thinking teaching programs. They concluded that successful instruction of critical thinking had been demonstrated and also that it argued against Piaget's stage theory.

In his chapter on critical thinking Hudgins (1978) described the studies he reviewed as having "generally been isolated in time and in intellectual origin from one another, and their cumulative impact is negligible. The sample sizes are small, and some of the reports are almost informal testimonials by a teacher of the values of instructing children to think critically. Perhaps the most crucial indictment to be leveled against this work is the failure by its authors to investigate the transfer effects of the instructional effects."

Curtis (1980) reported that there is a paucity of research studies which attempted to identify effective strategies for teaching critical thinking skills in social studies classes, and that those studies often produced inconsistent and inconclusive results.

In a study of implicit transfer of critical thinking in scientific situations, transfer did not manifest itself causing Dreyfus & Jungwirth (1980) to conclude that transfer of critical thinking from one context to another cannot be seen as probable.

Ponder & Davis (1982) reviewed the social studies research literature for the *Encyclopedia of Educational Research*, and concluded "the large body of research on critical thinking is ill-defined, and usually uses traditional teaching as a contrast, with traditional teaching left completely undefined (Wiley, 1977). As a result, inferences from this area are tendencies rather than conclusions. Marsh's review (1974) of 28 inquiry studies illustrates the problem of imprecision in studies of critical thinking in the social studies..." Eleven of the 28 studies favoured the inquiry method. Ponder & Davis reported however that Marsh discounted all 28 studies because of methodological inadequacies.

Ennis, Millman & Tomko (1983) reported finding eight (actually seven were reported) critical thinking instruction studies conducted since 1970 and indicated that there was no significant treatment effect in four, that the experimental group performed worse than the control group in two, and that the treatment was effective in one.

Cuban (1984) addressed the problems associated with the teaching of reasoning in the schools, including the organized inhospitality toward teaching of thinking existing in the schools, and recommended that research be conducted on promising ways to teach reasoning. Cuban noted that no clear agreement exists among cognitive psychologists on how to best teach children reasoning within specific content areas. Another problem Cuban raised is the thorny one of transfer of what is learned in one subject to a nonrelated subject. The issue of transfer is conspicuous by its absence in the literature.

McMillan (1986) reviewed the literature on the effects of methods, courses, programs, or overall college experience on critical thinking. Twenty-seven studies were located, nearly all old. Few effects were found for any of the treatments except overall college experience. McMillan concluded that lacking are a common definition, adequate instrumentation, and treatment articulation.

Since critical thinking has been broadened to include thinking, even learning, and intelligence, it may be instructive to include Snow's (1986) conclusion about the effects of educational
training to improve cognitive abilities. He observed that some broad interventions have had positive results. Some had initial improvements that waned. Some seemed to have positive effects on crystallized skills but negative effects on fluid skills.

In a recent revival of Raths’ teaching for thinking, Wassermann (1987) reported that Raths’ approach was effective in enhancing thinking in a number of studies, some of which were critical thinking studies, a generation ago.

Overview of the Effectiveness of Pre “New Wave” Thinking Skill Training Programs

The overviews and reviews of the pre “new wave” literature, some 12 in number reported herein, are hardly ringing endorsements of the effectiveness of the critical thinking skills training programs.

In 1983 Sternberg laid out eight necessary but not necessarily sufficient criteria for intellectual skills training programs:

1. based on independently empirically validated theory
2. theory should be socioculturally relevant to the subjects
3. explicit training in both executive and nonexecutive information processing
4. attention to subjects’ motivational needs
5. sensitivity to individual differences
6. linkage between the training and real-world behavior
7. careful empirical validations including transfer of training and
8. most importantly, modest claims.

Conditions Necessary for Success in Teaching Critical Thinking

This author is very concerned that the current critical thinking bandwagon will sputter to a stop before we learn how to teach critical thinking.

Five conditions need to obtain if the contemporary critical thinking/thinking bandwagon is to be able to teach students to think:

1. consensus on the definition of critical thinking/thinking
2. analyses of the psychometric soundness and factor structure of the developing “new wave” tests of critical thinking/thinking to determine the structure of critical thinking/thinking tests and inferentially critical thinking/thinking/per se, and to determine if the construct of critical thinking/thinking exists as a unique construct, separate from language ability, verbal ability, IQ
3. relation of the definition of critical thinking/thinking to the factor structure of the critical thinking tests, and vice versa
4. meta-analysis and synthesis of the effect sizes of the critical thinking/thinking instruction methods across the old literature to identify which instruction methods, if any, actually have been effective in the past
5. finally, and crucially, independent, methodologically sound, replicated, research on the effect sizes of the purported “new wave” critical thinking/thinking instruction methods to determine which instruction methods, if any, work under which conditions.

Consequences If Conditions Are Not Met

Following are the projected consequences if these conditions are not met. Each condition will be presented followed by its projected consequence if it is not met.

1. Consensus on the definition of critical thinking/thinking. If consensus does not obtain on the definition of critical thinking/thinking there will be a
plethora of different definitions with a correlative larger number of measures of critical thinking/thinking purported to be respectively appropriate. More importantly, there will be a correlative larger number of purported effective critical thinking/thinking instruction methods. This deluge of definitions, measures, and methods will result in conceptual chaos of such a magnitude as to preclude any systematic instruction of critical thinking/thinking. For example, without a definition of critical thinking/thinking we will be unable to even begin to determine when critical thinking/thinking has been successfully measured let alone successfully taught. See Follman (in press) for an elaboration of this definition problem. Shrag (1987) argued that too little thought has gone into determining what teaching thinking means. He also expressed concern about separating the thinking process from content and context. Finally he noted that he thought of the good thinker as one who exhibited a certain trait of character.

2. Analyses of the psychometric soundness and factor structure of the developing "new wave" tests of critical thinking/thinking to determine the structure of critical thinking/thinking tests and inferentially critical thinking/thinking per se, and to determine if the construct of critical thinking/thinking exists as a unique construct, separate from language ability, verbal ability, IQ. By psychometric soundness is meant adequate test-retest reliability as well as internal consistency reliability, and at least face validity, while efforts to establish concurrent, predictive, and construct validity are conducted. By factor structure is meant that items purportedly measuring the same thinking skill correlate higher with each other, than with items purportedly measuring other thinking skills, as well as associating together on the same factor. The importance of consistent factor structure across different data collection samples is that if items purportedly measuring the same thinking skill do not intercorrelate highly and also do not associate on the same factor, across the different samples, we will not know what the thinking skill is nor will we know what it is not. It is important to know what it is not because of the empirically well established correlation with verbal ability, language ability, scholastic aptitude and scholastic achievement (Follman, 1987). Another function of factor structure in tests composed of subtests is to break down the halo effect variance and sort it out among the different subtests. The problem with undifferentiated halo effect is that it is not clear what it represents. Without factor structure, i.e., separate subtest configurations, a thinking skill test could indicate an increase in pretest to posttest scores presumably as a consequence of some thinking training program, but it could not be determined what the thinking skill which was purportedly improved actually is.

It is not to be inferred that the necessary test construction and refinement implied in the paragraph above will be easy. Fortunately there is a sound psychometric base upon which to build, i.e., the 50 year pre "new wave" critical thinking test literature, as well as the venerable reasoning, problem solving, judgment, and other thinking test literature. See Stewart (1987) and Follman (in progress) for a treatment of critical thinking tests.

One model for the empirical examination of the critical thinking/thinking test factor structure would be to administer one or more "new wave" critical thinking/thinking tests as well as such old wave critical thinking/thinking subtests as recognition of assumptions, reliability—evaluation of evidence, whether or not conclusions follow, etc., along with such marker tests as deduction, induction, vocabulary, etc. These data would then be analyzed through correlational and factor analysis (probably principal components with both oblique and orthogonal rotations) to
determine if the a priori conceptual factor structure holds up a posteriori.

If the now developing critical thinking/thinking tests do not meet the conventional psychometric standards, especially reliability, we will be unable to determine the effects of methods of teaching critical thinking/thinking, as all other standards are too subjective for practical use.

If the now developing critical thinking/thinking tests do not develop stable factor structures it will be extremely difficult to determine what each test measures. Also, if the structure is either unclear, or clear but not disparate from language ability, verbal ability, IQ, there will be continuing confusion, ambiguity, and obfuscation in the definition of critical thinking/thinking. If this definitional confusion continues it will be difficult, if not impossible, to know what aspects of critical thinking/thinking to teach. It will be difficult to know what is being measured, critical thinking, thinking, language ability, some combination of the two or combinations of either critical thinking/thinking or language ability and some third, fourth, variable(s), etc. If we don’t know what the tests measure and/or if we can’t determine the effects of methods of instruction of critical thinking/thinking the current bandwagon will stagger to a stop. For an elaboration of the actual existence of critical thinking see Follman (in press).

3. Relation of the definition of critical thinking/thinking to the factor structure of the critical thinking/thinking tests, and vice versa.

If the structure of the critical thinking/thinking tests does not relate to the deductive definition of critical thinking/thinking there is little likelihood that the test scores will indicate that critical thinking/thinking has been taught, except randomly over a number of efforts. In order to clear up the conceptual chaos associated with the plethora of deductively determined definitions it may be necessary to work backwards, and determine the definition of critical thinking/thinking from the psychometrically determined factor structure of the tests, i.e., let the tail wag the dog, especially since no consensus has emerged among the deductively determined definitions, except to broaden the definition. This broadening of the definition can only obfuscate the definition, and the testing, and the teaching, of critical thinking/thinking, since there will be mismatches between and among the definitions, the tests, and the characteristics of the different critical thinking/thinking instruction methods.

4. Meta-analysis and synthesis of the effect sizes of the critical thinking/thinking instruction methods across the old literature to identify which instruction methods, if any, actually have been effective in the past. If we don’t use methods demonstrated effective we will also attempt untried methods which will produce problematic success at best. With little, or only problematic, success in teaching critical thinking/thinking, disillusion will soon set in.

5. Independent, methodologically sound, replicated, research on the effect sizes of the purported effective “new wave” critical thinking/thinking treatments to determine which instruction methods, if any, work under which conditions.

If independent, methodologically sound, replicated, research on the effect sizes of the purported effective “new wave” critical thinking/thinking treatments is not conducted, the bandwagon will roll on for a time, on the come, as it has thus far, until educators begin to question what gains are being made, and finding few, other than the ubiquitous testimonials, and also contemplating the diminishing returns associated with the cost and time of the treatments, disillusionment will set in, and we will have completed another down phase in the critical think-
ing/thinking cycle. There is developing evidence that this has begun as a recent survey of school administrators indicated that they felt that they needed more information on how to teach thinking and reasoning skills before adding thinking and reasoning skills instruction to their curricula (Education Week, May 27, 1987).

References


Educators, Houston, Texas.


Social Education, 26, 13-16.


Professor John Follman, College of Education, University of South Florida, Tampa, Florida 33620