PSYCHOMETRIC CONSENSUAL OPERATIONAL DEFINITION OF CRITICAL THINKING

John Follman

Introduction

In the last decade there has been a surge of interest in the venerable construct of critical thinking and critical thinking has become an education juggernaut. Ignored in this juggernaut has been the lack of consensus on the definition of critical thinking.

A main plank in the rationale for the current emphasis on enhancing critical thinking of children in general is the finding that 80% of third graders, more than half of seventh graders, and 36% of eleventh graders scored minimally or inadequately when reading critically on the National Assessment of Educational Progress (Vobejda, 1988, February 26). More specifically, less than 1% of third graders, 8% of seventh graders, and 23% of eleventh graders performed at the highest level on reading performance of 36,000 public and private school students.

Richard Paul, the chair of the National Council for Excellence in Critical Thinking, recently expressed his concern about the contemporary state of critical thinking: "With critical thinking swiftly becoming a marketable buzz-word that generates funding and profits, the educational community is well-advised to ask, 'Where's the beef?' before they buy in." According to Paul, there is an increasing number of superficial approaches to critical thinking being advocated, many of which have little or no substance (Critical Thinking, 1992,a). Further, in an independent piece in the same issue (Critical Thinking, 1992,b) the concern was raised that subjectivity is confused with reasoned evaluation and judgment by noting an incident in California in which a student essay was selected as an example of 'exceptional achievement' in reasoned evaluation, and that 150,000 copies had been disseminated nationally. Critical Thinking then observed that "the only thing it modelled was fluent subjective reaction."

Need for Definition of Critical Thinking

Three decades ago, Ennis (1963) noted that although critical thinking is generally recognized as one of the most important education goals, determination of the definition of the construct of critical thinking was needed. Additional evidence that this concern for a consensus on a definition of critical thinking is not ephemeral also adduces from the fact that the American Council on Education was seriously concerned, even earlier, forty years ago (Dressel & Mayhew, 1954).

Hart (1986) addressed the skepticism concerning critical thinking teaching programs and wrote that there was a lack of agreement or any substantive evidence to support a program for training in thinking skills. Nickerson (1989) noted that there are two compelling measures for testing for thinking, to determine how effectively thinking is taught, and to insure that instructional effort is being made on this objective.

A recent instance of concern about the contemporary state of critical thinking, specifically its definition, was indicated by Facione (1990) who convened a panel of critical thinking experts to try to develop consensus on the definition of critical thinking.

Components of Definition of Critical Thinking

Ennis (1962) articulated three critical thinking dimensions: logical, criterial, and pragmatic; and 12 aspects: grasping the meaning of a statement; judging if there is ambiguity in a line of reasoning; and judging contradictions, if a conclusion follows, if a statement is specific enough, if a statement is actually the application of a particular principle, if an observation statement is reliable, if an inductive conclusion is warranted, if the problem has been identified, if something is an assumption, if a definition is adequate, and if a statement made by an alleged authority is acceptable.

Beyer (1984) called for a clarification of the concept of critical thinking, specifically for clarification of exactly which skills constitute critical thinking. His suggestive final list consisted of: distinguishing between verifiable facts and value claims; determining the reliability of a claim or source; determining the accuracy of a statement; distinguishing between warranted or unwarranted claims; distinguishing between relevant and irrelevant information, claims, or reasons; detecting liars; identifying stated and unstated assumptions; identifying ambiguous or equivocal claims or arguments; recognizing logical inconsistencies in a line of reasoning; and determining the strength of an argument. These skills are found frequently in conventional critical reading tests.

For recent illustrations of the surfeit of critical thinking definitions see Beyer (1985) for an iteration of some two dozen definitions reported prior to 1985. More recently Follman (1991, October) compiled some 20 more definitions posited since Beyer (1985). Follman (1987) had expressed his concern that this surfeit of definitions, without a consensus on the definition, would...

Hart (1986) addressed the skepticism concerning critical thinking teaching programs and wrote that there was a lack of agreement or any substantive evidence to support a program for training in thinking skills. Nickerson (1989) noted that there are two compelling measures for testing for thinking, to determine how effectively thinking is taught, and to insure that instructional effort is being made on this objective.

A recent instance of concern about the contemporary state of critical thinking, specifically its definition, was indicated by Facione (1990) who convened a panel of critical thinking experts to try to develop consensus on the definition of critical thinking.

Components of Definition of Critical Thinking

Ennis (1962) articulated three critical thinking dimensions: logical, criterial, and pragmatic; and 12 aspects: grasping the meaning of a statement; judging if there is ambiguity in a line of reasoning; and judging contradictions, if a conclusion follows, if a statement is specific enough, if a statement is actually the application of a particular principle, if an observation statement is reliable, if an inductive conclusion is warranted, if the problem has been identified, if something is an assumption, if a definition is adequate, and if a statement made by an alleged authority is acceptable.

Beyer (1984) called for a clarification of the concept of critical thinking, specifically for clarification of exactly which skills constitute critical thinking. His suggestive final list consisted of: distinguishing between verifiable facts and value claims; determining the reliability of a claim or source; determining the accuracy of a statement; distinguishing between warranted or unwarranted claims; distinguishing between relevant and irrelevant information, claims, or reasons; detecting liars; identifying stated and unstated assumptions; identifying ambiguous or equivocal claims or arguments; recognizing logical inconsistencies in a line of reasoning; and determining the strength of an argument. These skills are found frequently in conventional critical reading tests.

For recent illustrations of the surfeit of critical thinking definitions see Beyer (1985) for an iteration of some two dozen definitions reported prior to 1985. More recently Follman (1991, October) compiled some 20 more definitions posited since Beyer (1985). Follman (1987) had expressed his concern that this surfeit of definitions, without a consensus on the definition, would...
result in this, latest, compelling cycle of critical thinking activity ending in a diverse, motley, incoherent, inchoate collection of assertions, recommendations, definitions, dispositions, concepts, constructs, tests, measures, etc., of critical thinking with no greater determination of the definition of the construct of critical thinking than existed prior to 1983. Since there is not consensus on what critical thinking is, there is not consensus on how to measure it. Since there is not consensus on how to measure critical thinking, it can not be determined if it has been taught. This condition has characterized critical thinking since its origin about 70 years ago. A parallel problem obtains with the construct of critical reading. For example, Bagford (1990) observed that there are two crucial reasons why teaching critical reading has not advanced over the years, no widely accepted definition of 'critical reading', and no acceptable tests to measure students' progress in critical reading.

Follman (1987) recommended that a practical approach to the determination of the definition of critical thinking is to use an empirical, operational, psychometric one. Specifically, all critical thinking tests should be collected in order to identify the tests and especially subtests appearing most frequently. Reviewing the critical thinking tests and subtests appearing most frequently is not novel having occurred at least three times in the literature. From such a consensus of the critical thinking tests, especially the subtests, a definition of critical thinking can be derived. Following are arguments in behalf of such a consensus.

Arguments for a Psychometrically Determined Operational Definition of Critical Thinking

The initial argument for this counting, frequency, empirical, inductive, operational approach, is its simplicity, just count.

The second argument is that it is a reasonable assumption that the critical thinking test makers over three generations gave considerable thought to the definition of critical thinking in the development of the specifications and items in their respective tests. It is also most likely that they were influenced by the specifications and substance of prior critical thinking tests reaching as far back as Watson's (1925) Test of Fair Mindedness. This building on the shoulders of their predecessors represents a developing, continuing consensus, perhaps the only significant continuous consensus.

A third argument in behalf of the empirical counting, frequency of subtests determination of the definition of critical thinking is by analogy based on such taxonomies of intellectual skills as Bloom's Taxonomy and Guilford's Structure of the Intellect. Bloom's Taxonomy includes knowledge, comprehension, application, analysis, synthesis, and evaluation. Guilford's model includes units, classes, relations, systems, transformations, and implications. By analogy the consensus of critical thinking subtests can serve as an appropriate taxonomy. Also see Norris & Ennis (1989) for a compendium of topics re evaluating critical thinking.

A fourth argument in behalf of the inductive, critical thinking test and subtests' consensual operational approach is that of default. While there is no metaphysical certitude associated with either the inductive approach or with the deductive approach, it can not be gainsaid that the deductive approach has not succeeded in determining an agreed upon definition.

For additional evidence of the importance of the instruments of critical thinking in the historical development, articulation, and delineation of the construct of critical thinking see Follman (in progress) for a massive compilation in excess of 700 critical thinking correlations representing the critical thinking nomothetic network correlation knowledge base. It will be noted therein that nearly all these correlations are instrument oriented.

Some additional evidence of the intrinsic importance of the critical thinking tests to the construct of critical thinking is in a recent meta-analysis of the efficacy of explicit efforts to teach critical thinking where of the 20 "studies considered suitable" 15 used only two tests, 10 used the Watson-Claser Critical Thinking Appraisal and five used the Cornell Critical Thinking Test (Bangert-Drowns & Bankert, 1990, April). Critical Thinking Tests/Subtests Consensuses

The remainder of this paper is a review of the literature of consensuses of critical thinking tests/subtests.

A comprehensive collection of critical thinking components was compiled by Herber (1959). Herber taxonomized some 41 critical thinking components into five main categories; basic attributes, selection of a problem, gathering evidence, and analysis of data, and establishing conclusions.

An early instance where a consensus of critical thinking subtests was developed was a dissertation by Handel (1965). The subtests (and corresponding frequency rank orders) appearing most frequently were: author's bias (10), generalizations (8), inferences (8), relevancy (7), assumption (4), logic (6), reliability of source (5), fact vs opinion (4), conclusions (4).

The next subtests' consensus, of 37 works, was also a dissertation, by Lanseigne-Case (1967) who conducted a useful review of the then current critical thinking tests. The subtests (and frequencies) appearing most frequently were: author's bias (18), relevancy (15), fact vs opinion (14), inferences (13), generalizations...
(12), reliability of source (12), assumptions (9), conclusions (9), and logic (7). Lanseigne-Case also developed a measure of detection of author's bias and also inference for sixth graders.

The fourth consensus, also a dissertation, was by Curry (1971), who conducted this consensus in order to determine the components in the construction of a critical thinking test. Curry reported the following subtests (frequencies): reliability of source (18), fact vs opinion (8), data (15), assumption (13), logic (13), author's bias (11), relevancy (11), conclusions (10), generalizations (8), inferences (7), and propaganda (7).

These four researchers have exhausted the then contemporary literature of authors' critical thinking opinions, theories, skills, abilities, aspects, facets, especially tests and subtests, and their frequencies of occurrence.

Another, much more recent, compilation of some 25 critical thinking tests was collected by Stewart (1979) who reported that five measures tested a large number of critical thinking abilities. These five tests were: Cornell Critical Thinking Test Level X; Cornell Critical Thinking Test Level Z; Logical Reasoning Test; Pearson-Hyram Test of Critical Thinking; Watson-Glaser Critical Thinking Appraisal. Baker (1981) also compiled the contemporary critical thinking tests. Also see Follman (in progress, b) the the most recent inventory of the critical thinking tests.

Gubbins (1985) developed a useful Matrix of Thinking Skills. Also anent thinking, the overlap between critical thinking components and the new GRE Analytical Test was noted by Duran, Powers & Swinton (1987, April).

A more recent taxonomy of 43 examples of critical thinking types encompassing 19 relevant references from 1946 through 1985 was prepared by Dick (1987). Dick classified these 43 types of critical thinking into 15 superordinate types. Dick then categorized these 15 superordinate critical thinking types into five major types of critical thinking: identifying arguments, analyzing arguments, evaluation—assessing claims, assessing arguments, inference—querying evidence, conjecturing alternatives, drawing conclusions, explanation—stating results, justifying procedures, presenting arguments, self-regulation—self-examination, self-correction.

Note that Paul's (Critical Thinking, 1992, a) concern about critical thinking came after Facione's (1990) implied Delphic closure on the issue of the definition of critical thinking.

**Consensus of Censuses**

The consensus of critical thinking subtests across these censuses follows.

A critical issue in establishing any consensus is the level of explanation desired, how molar or how molecular should the explanation be.

At the molar, “test” level analysis, main components (Herber, 1959), main dimensions (Facione, 1990), main types (Dick, 19xx) the semantic inconsistencies in category labeling allow only a crude, tentative taxonomy, to wit analysis, inference.

At the molecular, "subtest" level analysis, the frequencies indicate the following frequency rank order of subtests: author's bias; reliability of source; inference, logic; generalizations; fact vs opinion; relevancy; conclusions; assumptions, etc. It is noted that a more specific, more molecular taxonomy is precluded here because of the known but undifferentiatable overlap in categories in Handel (1965), Lanseigne-Case (1967), and Curry (1971) because all the categories in all the studies they cited are not explicitly enumerated by the respective study from which they were obtained. It is considered that the results at this level of analysis are tuitional.

As noted previously, Ennis (1963) three decades ago and Dressel & Mayhew four decades ago noted that determination of the definition of the construct of critical thinking was needed. Also calling for a definition of critical thinking have been Beyer (1984), Hart (1986), Follman (1987) and Facione (1990). An essay in this direction has now been accomplished through the empirical, counting, frequency approach above to suggest the critical thinking tests and subtests identified herein as the consensus definition of critical thinking.

*(for references see page 46)*

John Follman teaches in the Department of Psychological and Social Foundations at the University of South Florida at Tampa.