

# LESS THAN CRITICAL THINKING

---

***BY ROBERT W. TUCKER***

***ORIGINALLY PUBLISHED – 1996  
ASSESSMENT AND ACCOUNTABILITY FORUM***

### **EDITOR'S NOTE**

This article on critical thinking (CT) examines CT as a concept and its evolution as a construct in the behavioral sciences. Conceptual and empirical problems regarding the current generation of CT tests are briefly reviewed and alternatives to current CT tests re examined.

--Editors, *Assessment and Accountability Forum*.

### **CONTACT INFORMATION**

16 Crowley Lane, PO 1230  
McCall, ID 83638

520/ 908-7815

800/516-9571

[www.InterEd.com](http://www.InterEd.com)

# Assessment and Accountability Forum

Less Than Critical Thinking

## Introduction

We care more about critical thinking these days--having the skills ourselves and developing them in our children. Significant drivers of the increased attention to critical thinking--such as the near-polynomial growth of knowledge, the transformation to an information-based society, and the central role leveraged knowledge now plays in gaining economic advantage--are all tied to one's ability to critically evaluate (i.e., comprehend, organize, assess, and adjudicate in relation to a goal) and apply only that minuscule part of the total knowledge that will do the most good for the problem at hand. Once relatively easy if you were reasonably smart, this is now a tough task for the best of minds. In education circles, Critical thinking (CT) gained widespread recognition as a behavioral science construct in the 1980's when Goodwin Watson and Edward Glaser's *Critical Thinking Appraisal* became a widely used tool for assessing the effects of undergraduate education on reasoning skills. The publication of *A Nation at Risk* in 1983 fueled the fires for increased emphasis on certain core skills, including critical thinking.

In 1988 and 1989, 46 experts participated in a project designed to reach a consensus on the core elements of CT for college graduates by using the Delphi technique. In the same year, the nation's governors adopted the National Educational Goals that identified CT as a core skill. By 1990, the SCANS project (Secretary's Commission to Achieve Necessary Skills) and the publication of *America's Choice: High Skills or Low Wages* had focused the nation's attention on the importance of core skills, including critical thinking, to maintaining national productivity and global competitiveness. The National Educational Goals were adopted under Goals 2000 and became law in 1993. In November of 1991, the U.S. Department of Education sponsored a workshop on *Assessing Higher Order Thinking and Communication Skills in College Graduates*. Among the participants were the most productive thinkers on the subject. Largely through the National Center for Education Statistics, the Department of Education contributes to scientific and public understanding about methods for assessing and teaching CT in ways that are congruent with the Department of Labor's school-to-work transition initiatives.

The CT agenda has more than political and governmental backing. Professional education programs, nursing education for one, have made the assessment of CT skills a top priority. DeVry Education plans to assess the development of CT skills in its technical programs. We developed CT assessment components for the University of Phoenix as early as 1986. Increasingly, human resource development and corporate education departments are asked to find better ways to assess and improve thinking and decision skills. As a nation, we are weaving CT skills into economic models of business productivity. In the United States, intellectual systems and processes create 79 percent of all jobs and 76 percent of the GNP. While content knowledge is a crucial part of this value creation, CT skills are the adjudicative engines that drive everything from strategic boardroom decisions to the creative responses of

# Assessment and Accountability Forum

## Less Than Critical Thinking

a software company's help desk. CT skills are a hot commodity and U.S. corporations are finding ways to leverage them through the application of creative organizational structures and information systems designed to replicate knowledge and thinking units within their organizations.

### ***What Is There to Mean by Critical Thinking?***

Given the obvious importance of CT skills, why is it that so many studies and essays on critical thinking begin with a tedious definition of the subject? How much do we know about CT and how does one decide that it is taking place? Why is there no dominant theoretical model of CT to generate empirically testable propositions? Why are some definitions of CT peppered with marginally related social prescriptions? Why is it that so much talk about CT centers on Bloom's taxonomy of the cognitive domain--a once sensible model that has failed its empirical tests and should be discarded? Is CT a collection of skills and attributes the exact composition of which is to be legislated by experts through a consensus-building exercise? Is there any empirical finding that would cause one to reexamine the criteria for, or definition of, CT? In short, will the behaviors we decide to call evidence of CT be determined more by stipulation and social convention, or by empirical study?

Was Albert Einstein a critical thinker? Is Julia Child? Can anyone who achieves the position of President of the United States be presumed to be good at CT? Is Mother Theresa who she is, in part, because she is a skillful critical thinker? Can one be good at critical thinking in one walk of life and inept at it in another? Can one 'will' oneself to be a critical thinker about a specific problem and fail to exercise the necessary will in a second very similar context? Can we call it CT when someone uses his rhetorical, logical and analytic skills to achieve a preconceived (and possibly, non-rational) end? Can one become a better critical thinker through practice? Is there a set of cognitive and/or affective tools one can acquire and use to function as a critical thinker? Are there necessary conditions to critical thinking skills (e.g., rhetorical and informal logic skills) which are not themselves the skills of critical thinking; if so, what are the sufficient skills to critical thinking? Does it mean the same thing to be a critical thinker in banking as when writing poetry? Can one be a critical thinker when writing poetry? Is a literary critic more likely to be a critical thinker than a writer? Is there a set of critical thinking skills; or is the notion of critical thinking a generalized idea that has different meaning in different contexts?

The flood of ideas triggered when contemplating the above questions characterizes the complexity and uncertainty that surround ordinary language concept. Despite the contributions of the philosophers and behavioral scientists who worked on the Delphi and other projects, there is still much inconsistency, incongruence and fundamental

# Assessment and Accountability Forum

## Less Than Critical Thinking

incommensurability of perspective when we talk about critical thinking. We should exploit this messiness more and should not in any way suppress it, including the de facto suppression that occurs when a concept is defined by fiat or prematurely shoe-horned into a consensus model.

It would be helpful to have the results of a few dozen independently performed analyses on the concept of *critical thinking*, especially if some of them were performed by business and the professional leaders. The main purpose of conceptual analysis is to discern, organize and formalize whatever collective wisdom is embodied in the linguistic conventions of naming and categorizing the instances and generalizations of our observations. Conceptual analysis works from cases to identify common elements via a simple process. It begins by identifying a wide range of cases about which we can say with certainty, “these are instances of critical thinking” (call these “positive cases”). Next comes the identification of a wide range of cases about which we can say with equal certainty, “these are instances of thinking that is *not* critical thinking” (call these “contrary cases”). Last is the identification of a wide range of cases that seem like CT but cannot be classified with certainty (call these “borderline cases”). There should be at least ten good cases in each of the three categories. The next step is to extract the common elements from the positive cases. For example, a common element in all examples of CT might be that the thinker considers at least three alternative interpretations of a set of purported facts and tests the consequences of each. The same process is repeated for the contrary cases. If the conceptual analysis turns out well, the reward will be a short list of criteria that must be met to call something “critical thinking.” If these criteria are well-formed, they will reduce the uncertainty in classifying the borderline cases. As obvious as the outcome of this exercise might seem, our definitions and assertions about CT will be aided if all who aspire to be CT scholars will, first, gather a small team of thinkers and spend a few days or weeks performing a thorough conceptual analysis. (If you do such an analysis, I would be pleased to hear from you.)

Conceptual analysis has limitations. Some complex concepts have no common elements to all of their cases. (Consider the concept ‘game’ where one soon realizes that there are no non-trivial elements in common to all things that we call games.) We say that such ideas are *family resemblance concepts*, after the work of Wittgenstein. Individual elements of family resemblance concepts are to the concept itself as individual strands of a rope are to the full length of the rope; no one strand runs from one end of the rope to the other, yet there is a single piece of rope for all to see. Conceptual analysis has many other limitations, including the fact that it provides analytic information and does little to advance the creation of a prescriptive model based on empirical findings not embedded in common language.

## ***How Critical Is Our Thinking on CT?***

How much do we know about CT as an ordinary language concept? How much of this knowledge has been carried forward to create a theoretical construct open to the kind of scientific examination that will satisfy us that the phenomena we want to call CT are being examined? It depends. First we have the knowledge that comprises our ordinary language understanding of CT. This is received-view knowledge, gloss, intelligence buried in our language, and folk psychology. As I suggested above, one gets at this collective intelligence best through conceptual analysis combined with the judgment of experts. This knowledge, from all of its sources, is both massive and massively unfocused. CT, as an ordinary language concept, cannot be distinguished from related ideas, such as IQ on one side, problem-solving and decision-making on another, and cleverness and cunning on yet another. We seldom say, “He is smart but not a good critical thinker” although we do say, “He is very intelligent but can’t make a good practical decision.”

Second is the knowledge we should have from organizing the concept of CT into a theory or model that possesses scientific merit. We will know that we have advanced to this stage when we have decent answers to the following questions. Does the model of CT describe or subsume important phenomena? Does it do so in an exclusive way (e.g., not isomorphic with definitions of IQ)? Is the model scrutable? Does it produce falsifiable propositions? Does it define adequate metrics? Is it conceptually congruent (i.e., doesn’t produce too many contradictory propositions)? Right now, the concept of CT is a loose collection of ideas that have not been sufficiently developed to warrant advancement to the status of a well-formed, empirically grounded construct.

Last is the kind of knowledge that we should derive when a theory or model is employed to gather research to produce a substantial body of empirical evidence. To pass this test, it is not so important that the measures be unique (redundancy can be valuable) but that they be valid. Do certain kinds of CT scores lead to valid understandings of human performance and do those understandings add to the relevant knowledge base in a unique way? If an individual’s CT subtest shows superior skills in drawing inferences, does that person actually possess superior inferential skills as measured by several independent methods? Does the “Inference” subtest correlate highly with other measures of inferential skill? Most important, does any scale, subtest or combination of subtests lay a foundation for understanding, explaining, or predicting things that are currently not well understood, explained or predicted? Today, the links between models and measures of CT are weak.

# Assessment and Accountability Forum

Less Than Critical Thinking

## ***Metrics: Too Clever by Half***

We can seek and achieve useful ends under the guidance of a bad conceptual model. To realize this, one need only recall that it was possible to predict planetary motion under the Ptolemaic model of the solar system; or that, prior to the definition of an acid as a proton donor (which had to await the discovery of protons), one could classify acids with reasonable accuracy using ordinary litmus paper. Closer to home, the complex business of predicting success in academic studies was accomplished prior to the creation of the IQ tests. One could argue that this reasoning applies to the concept of CT: never mind the theoretical issues (prescribe them away if necessary), get on with the measures.

This is a serious mistake. Ideas evolve by showing stable and predictable relationships with other ideas and constructs, and with empirical phenomena. From these relationships we gain the ability to explain, understand, and predict. Today, CT is an ill-defined human attribute looking for something to explain, understand and predict. It is also a collection of “wouldn’t it be nice if CT were this’s”--i.e., moral prescriptions that may have deep functional value but do not strengthen the concept of CT. For example, I hope that all critical thinkers will possess the humility that comes with a mature intellectual perspective. I also believe that more critical thinking will get done, in the long run, if one is intellectually humble. But it is a far cry from these hopes and beliefs to asserting that intellectual humility is an attribute of CT. If this assertion is intended to be a criterion of CT (i.e., a definition), there remains the very large question of who gets to decide all about intellectual humility and what its inclusion would do for the concept of CT. If the assertion is intended as a generalization from extant empirical evidence, where is the non-tautological evidence that the humble do better at critical thinking? I can see only one tenable position for those who would like to have intellectual humility be a part of CT as a scientific construct. They must conduct empirical studies that examine and describe the relationship of intellectual humility to whatever else we think makes up CT. Whatever the outcome of these studies, we will have advanced our thinking on CT--intellectual humility will be in, out or a definite maybe. The same kind of examination needs to take place with respect to other reasonable assertions about the meaning of CT.

The above studies are crucial to the study of CT. We need much better organizing constructs if CT is to become a theoretical model capable of generating empirically testable propositions of sufficient quality and number. For the sake of the point, imagine that we discover critical thinking attributes CT1 ...CTn to be determined (i.e., essentially predicted) by a specific genetic configuration. Otherwise very intelligent individuals (e.g., with respect to memory, eidetic imagery, computational skills, etc.) who do not possess this genetic configuration are unable to do CT-like things such as anticipate consequences, weigh alternative courses of action, cases, and so on. Assume that the presence or absence of this genetic profile correlates highly with having (or not having) six of the eight attributes most

# Assessment and Accountability Forum

## Less Than Critical Thinking

commonly thought of as CT. Would this finding result in a substantial modification of the current construct of CT, as was the concept of an acid reformulated when particle physics provided a better theoretical model? The answer is most likely “yes.” While we may or may not find a genetic explanation for the specific kind of thinking we want to call *critical* thinking, we do need a better way than we have now to organize our knowledge, values and observations so that CT gains both explanatory and predictive power. When the concept of CT is reformulated, it will affect current classifications--some included cases will be excluded and vice versa.

### **Limitations of Current Instruments**

No tests of CT have been demonstrated to function as more than weak-to-moderate proxy measures for IQ and/or achievement tests. To date, the descriptions, explanations and predictions that can be made with CT tests can be made better with other instruments. Given the nature of the associations between CT tests, and IQ and achievement tests, it would not be surprising to find that tests like Miller’s Analogies would measure what current generation CT tests purport to measure, and may exceed CT tests with respect to predicting the outcome of CT-like tasks.

The greatest limitation of the current generation of CT instruments is that they do not possess an underlying empirical structure necessary to support the instrument’s *a priori* definitions of subtests and the constructs that the subtests define. For example, the most widely used instruments each have a subtest that purports to measure the ability to make sound inferences. (Terminology varies and “sound” is to be taken as conforming to applicable canons of reasoning.) These subtests have been created by combining several questions designed to test inferential skills into a single variable which is then treated as an instantiation of the construct, “inference making.” The tests’ creators offer two bases for believing that these questions measure inferential skills. First, inferential skills are required to answer the questions correctly--this is called “face validity.” Second and presumably, experts in philosophy, cognitive science and education think that these questions will measure inferential skills--this is called “content validity.” When students’ scores on each subtest are reported, they are usually in the form of raw scores, which reflect the sum, the mean or some other linear arithmetical representation of the combined scores of the “inference” questions. Percentile rank or other norm adjusted scores may also be presented, or can be looked up in a table, indexed to the raw score.

This procedure for constructing test scales is generally regarded as acceptable--as far as it goes--but it is the first and easiest step in a series of steps that ultimately lay a foundation for the claim that one is measuring a specific intellectual or attitudinal attribute. These additional

# Assessment and Accountability Forum

## Less Than Critical Thinking

steps probably should have been performed on the current generation of instruments before the assessment was released, at least as a commercial tool. Unfortunately, when exploratory and confirmatory factor analytic techniques are performed on data sets derived from the current generation of CT tests, one finds weak global factors that do not correspond to the subtest definitions. These findings are more than a statistical curiosity of interest only to measurement scientists. They mean that a subtest of, say, inferential skills, will describe neither a consistent nor meaningful relationship, across cases, to the skills measured by its individual questions. It may also mean that the individual questions themselves are not discriminating among levels of performance on any intended measurement dimension. Without solid, empirically derived subtests, it is not possible to establish discriminant validity for these subtests, so the validation process stops here--far short of the goal.

One should not leave this area without issuing a warning about the potential misuse of reliability statistics. The reliability of a test is a necessary property, but one that is not especially difficult to achieve. Almost any well constructed test will achieve acceptable reliability statistics. Two reliability statistics confused test validity are the Kuder-Richardson and the alpha coefficient (Cronbach's alpha). These statistics have only an indirect relationship to validity, even though test manuals may report the alpha coefficient of a scale in a way that might be interpreted as evidence of the scale's convergent validity. Alpha coefficients are sensitive to the number of items in the scale. A large group of similar questions will generally produce a satisfactory coefficient statistic regardless of the exact meaning of the questions or the underlying structure.

For years, CT tests have been administered in a pre/post design intended to assess improvement in general reasoning skills caused by undergraduate education. Do data from the current generation of tests provide a rational basis for this action? To the extent that CT tests correlate highly with intelligence tests or subtests, there is not much justification for employing them in a pre/post fashion to assess learning outcomes of undergraduate education, although they might be used as one component of an admissions assessment battery. Since the single factor product of some of these tests also correlates moderately well with some achievement test scores, there is a weak rational basis for using CT tests for pre/post assessment of undergraduate learning outcomes--although not necessarily as a CT assessment. It is also possible, perhaps likely, that the moderate global correlations between CT tests and other tests are best explained as a cross product of an ability factor or factors that underlie both kinds of tests. One would be well advised to consider available alternatives before basing important curricular or programmatic decisions on the outcomes of CT tests.

# Assessment and Accountability Forum

Less Than Critical Thinking

## Next Generation CT Tests

From an examination of the way we talk about and classify thinking, it seems that there is a kind (or kinds) of thinking that we are inclined to call *critical thinking*. We use this term just because the thinking we call *critical thinking* is somehow distinguishable, even if it is not separable, from other kinds of thinking, even otherwise very intelligent thinking. So much for what we know and agree upon.

From this modest center of agreement, it does not follow that any professional or ideological group means the same thing when they call something *critical thinking*. In fact, there appear to be great differences in the way the term *critical thinking* is used in ordinary communication and the way it has been defined in the work of scholars. These differences arise from the way each group has approached the issue. In the larger world of commerce, the phenomena to which the term *critical thinking* is ascribed are actually constellations of goal-directed behaviors which take their meaning in specific contexts. The use of the term in these contexts may be with considerably less linguistic precision than would satisfy the scholars but it does not mean that these uses have less predictive value for the behavior of interest. On the other hand, most scholarly work on *critical thinking* amounts to a debate about lists of attributes which might qualify as criteria or definitions of the term. Scholars have employed a variety of strategies to arrive at these definitions. Only a very few of these strategies have involved systematic extrapolation from actual instances of thinking. The predominant strategy has been to derive the definition of *critical thinking* as a subset of what philosophers and educators think about thinking. Having been trained as a philosopher, I have high regard for the value of their work. However, I do not believe that the views of these scholars should be over represented in the definition of a term that holds such broad social meaning. Most of these educators and philosophers have regarded *critical thinking* as a set of skills and dispositions that can be applied to a variety of subjects and problems. In other words, they believe that *critical thinking* is a generalizable attribute, more-or-less free of the specific contexts in which it is manifested. This view may be correct, or at least partially so, but it seems more likely that the kind of thinking that the nursing professions wants to call *critical thinking* has little in common with the kind of thinking to which the Longshoreman's Union wants to apply the same appellation.

In my experience, context can modify *critical thinking* processes, including canons of reasoning, in non-trivial ways. This does not mean that context can modify the rules of informal logic (although it sometimes does, as when a community of concept sharers deliberately eschews a rational convention). It does mean that one's mastery of rational processes in one discipline does not ensure one of appearing to be (or being) rational in another discipline. The problem is more than one of learning new terminology. It is one of learning shared values and practices, language games, strategy and precedence for the tools of informal logic and, especially, of mastering the list of included and excluded concepts. The

# Assessment and Accountability Forum

## Less Than Critical Thinking

combinations of factors unique to a specific context can create non-trivial exemplars of *critical thinking*. Consider the following environmental variables which might affect the meaning of *critical thinking* in situ.

<b>Factors Which Might Affect a Discipline's Definition of Critical Thinking</b>		
<i>Factor</i>	<i>Polar Attributes<sup>1</sup></i>	
Primary Goals of Discipline or Profession	Application of Knowledge	Creation of Knowledge
Moral and Other Normative Ground in Which Critical Thinking Is Manifested	Low Moral Content or Implication, Few Applicable Ethics	High Moral Content, Well Defined and Applicable Ethical System
External Laws and Regulations Governing Discipline or Profession	Low, Flexible, Circumscriptive	High, Rigid, Pre- and Proscriptive
Source of Knowledge in Discipline or Profession	Normative, High Social Component	Objective, Lower Social Component
Precision of Knowledge Base	Low, Diffused	High, Focused
Upside Potential Vs. Downside Risk of Application of Critical Thinking Within Discipline or Profession	High Downside; Low Upside	High Upside; Low Downside
Formal and Defacto Structures of Discipline or Profession	Formal, Command, Complex	Informal, Market, Flexible
Relations Between Thinking and Useful Outcomes	Linear or Well Defined Exponentiation, High Predictability	Non-linear, Step Function, Quantum, Low Predictability
<sup>1</sup> These dimensions are illustrative only.		

A review of the above table in terms of the *critical thinking* practiced by a trauma center nurse as contrasted with the *critical thinking* practiced by an education theorist produces some interesting contrasts of relevance to the definition of critical thinking. For the nurse, critical thinking must remain within a well defined matrix comprised of ethical, scientific, technological, professional, and rule and policy considerations. Thinking beyond these boundaries (or at least too far beyond them) might be *creative*, but not *critical*, unless the concept of *critical thinking* is to be stripped of its utility function. For the education theorist, most of the nurse's boundaries do not exist and thinking outside the boundaries that do exist (e.g., canons of reasoning) may be precisely what is needed to improve a piece of theoretical

# Assessment and Accountability Forum

## Less Than Critical Thinking

landscape. Based on our research to date, I do not see a high degree of commensurability between the kind of thought that is deservedly called *critical thinking* in such diverse disciplines. I do not doubt that there are trivial and perhaps non-trivial common elements between such diverse disciplines but their nature needs to be explored a discipline at a time.

### ***A Developmental Approach***

The above issues of definition are important because they lay the foundation for conducting studies to answer our questions about *critical thinking*--and there are many questions. Do we want to define *critical thinking* in terms of its stable constitutional (e.g., genetic and early environmental) components, or do we want our definitions to incorporate or even emphasize those attributes that we believe can be successfully taught? Is it more important that a definition of *critical thinking* lead to useful measures and predictions of behavior or that it be congruent with modern theories of intellect? At present, there are many competing definitions of *critical thinking* with scant reason to choose one over the other. The National League for Nursing has recognized this complexity when it both required all NLN accredited schools to assess *critical thinking* but left the definition of the term up to the determination of the individual schools. The current generation of tools designed to assess *critical thinking* do not address most of the questions raised above or in Part I of this article. On balance, they are the result of an imposed definition which has not benefitted from a serious discussion of its practical or theoretical consequences.

Many of the unresolved issues in creating a broadly useful model for assessing and teaching *critical thinking* are unlikely to be resolved until there is considerably more empirical evidence about what *critical thinking* means in various professional contexts and what components of this meaning are responsive to educational intervention. For this reason, I advocate a predominately pragmatic and experimental approach to the development of assessment tools. Right now, the formal assessment of *critical thinking* is concentrated in a few educational areas. Most of these areas fall within the humanities and several areas of professional education. The National League for Nursing requires that students' *critical thinking* skills be regularly assessed as a condition of the school's accreditation. Most regional accrediting bodies strongly urge such assessments. Our conversations with other professional accrediting bodies suggest that the formal assessment of *critical thinking* will soon be added in a number of other professional disciplines. At the same time, most of these professional bodies have expressed concern about the practical constraints of conducting assessments in today's college environment. Students are already over assessed (or believe they are) and examinees' motivations to do well on the assessment can no longer be taken for granted.

# Assessment and Accountability Forum

## Less Than Critical Thinking

Based on the practical constraints and opportunities presented in today's environment for the assessment of *critical thinking*, paying attention to the following considerations may increase the probability of producing a useful assessment tool.

### Meet Environmental Constraints

If one can secure 8 to 10 hours of assessment time, a group of highly motivated examinees, and a healthy budget, a competent evaluator can assess *critical thinking* to meet the needs of a specific professional context. For the rest of the world, one can secure one hour of examination time, marginally motivated examinees, and a budget that is suited only to mass standardized testing. Most useful tests of *critical thinking* will be inexpensive, interesting to take, and will not require more than 45-50 minutes of an examinee's time.

### Employ Efficient Questioning Methods

At present, there is probably no more efficient question model for the large scale assessment of *critical thinking* than Michael Scriven's multiple-ranking question. If designed well, questions written to this design can assess a minimum of two dimensions simultaneously and produce four to six data points in very little more time than required to produce a single data point using the standard multiple-choice format. An assessment tool recently developed by InterEd for the assessment of *critical thinking* in nursing gathers 170 data points with a 45 minute administration time using Scriven's multiple-ranking technique. To realize further efficiency gains, this instrument simultaneously assesses for the examinee's ability to rank principles along a selected health care dimension and for the examinee's selection of dimension(s) to rank.

### Assess in Context

Since the prevailing view is that *critical thinking* is (or should be) context free, some evaluators will object to the inclusion of context as a basic element of *critical thinking*. Given that current evidence is insufficient to resolve the many questions about the generalizability of *critical thinking* skills, the best investigative strategy for now is to create assessments in specific professional contexts. The downside risk of introducing context into this *critical thinking* model is minimal. If context does not introduce non-trivial changes into the nature of *critical thinking*, then the effort will have gone mostly into ensuring that the language is suitable and no empirical support will be found for the 'context' dimension.

### Assess at the Critical Level

Current assessment tools allocate virtually all of their measures to the assessment of skills which are necessary to *critical thinking* (e.g., drawing inferences, generalizing) but may not be sufficient for it. This approach assumes that having what might be called foundational skills to *critical thinking* will ensure that the skills can be contextualized and resolved. This is a doubtful assumption. Conversely, a correct resolution of a contextualized case-oriented problem does seem to require that the foundational skills be in place. This logic makes

# Assessment and Accountability Forum

## Less Than Critical Thinking

possible an alternative of assessing skills which intellectually subsume (and require) these lower order skills but do not assess them directly. As one example, examinees can be asked to rank a half dozen principles as they apply to a specific case in which a decision is called for. The case and principles can be constructed so as to require such skills as *recognition of assumptions*, *evaluation of arguments*, etc. but the act of ranking principles requires more than any of these lower order skills--it requires a synthesis and adjudication of all of the lower order information. A potential downside of evaluations that subsume lower order evaluations is that the scores may be diagnostically weak. A failure to recognize an assumption and a failure to correctly evaluate an argument may produce the same kind of error in ranking principles and thus the evaluation would not point to a useful intervention. It is also possible that the *pattern* of errors in ranking principles will provide hints that are diagnostic to lower order skills. An examinee may display a pattern of errors in ranking moral principles that is not displayed in ranking similar non-moral principles. We are examining this issue right now but have no concrete results. Test developers and users will have to weigh these risks against the benefits this strategy may produce.

### Focus on Core Issues

The sketchy state-of-theory and the unfocused empirical findings regarding *critical thinking* create a narrow set of core issues to assess. While there are many who wish to assess to every taxonomic element produced by contemporary scholarship in *critical thinking*, it would serve us well to note the lack of empirical support for this approach. Factor analysis of data sets produced by instruments claiming to assess five or more *critical thinking* skill sets do not provide empirical support that these skill sets either exist or are being measured by the instrument.

In the interim, I propose a slim three-element model as the most that can be justified at this point. The elements of this model are: *agency*, *adequacy*, and *context*. Agency describes the axiological, teleological and dispositional (including habitual) elements of *critical thinking*. Agency is assessed, for example, when identifying the value processes associated with identifying bias in sources. In its idealized state, *agency* is the disposition to think critically, regardless of how well one might do it. Adequacy describes the fidelity of the *critical thinking* process to applicable standards. Measures of adequacy assess the learnable skills which overlay and give contextual meaning to the largely fixed intellectual skills that are foundational to rational processes. In its idealized sense, *adequacy* is the intellectual capability to think critically. Context describes the elements of *critical thinking* (including adaptations of canons of reasoning) which are defined and given specific and actionable meaning, or assigned higher or lower precedence (significance, salience, etc.) by the specific *critical thinking* environment. Context defines the necessary conditions to the correct resolution of a practical case.

# Assessment and Accountability Forum

Less Than Critical Thinking

The assessment of *critical thinking* in professional contexts is a policy priority for the Phoenix Institute and a top product objective for InterEd. We will keep readers of *Assessment and Accountability Forum* informed of progress and would like to hear from you if you are doing work on the assessment of *critical thinking* or have comments on what you have seen here. We are especially interested in hearing from those who have assessed *critical thinking* (by whatever means) to produce useful information. Please send e-mail to [rob@InterEd.com](mailto:rob@InterEd.com).