The Disposition Toward Critical Thinking: Its Character, Measurement, and Relationship to Critical Thinking Skill

PETER A. FACIONE *Santa Clara University*

Abstract: Theorists have hypothesized that skill in critical thinking is positively correlated with the consistent internal motivation to think and that specific critical thinking skills are matched with specific critical thinking dispositions. If true, these assumptions suggest that a skill-focused curriculum would lead persons to be both willing and able to think. This essay presents a research-based expert consensus definition of critical thinking, argues that human dispositions are neither hidden nor unknowable, describes a scientific process of developing conventional testing tools to measure cognitive skills and human dispositions, and summarizes recent empirical research findings that explore the possible relationship of critical thinking skill and the consistent internal motivation, or disposition, to use that skill. Empirical studies indicate that for all practical purposes the hypothesized correlations are not evident. It would appear that effective teaching must include strategies for building intellectual character rather than relying exclusively on strengthening cognitive skills

Résumé: Les théoriciens ont supposé qu'il y a une corrélation positive entre les habiletés de la pensée critique et la motivation constante à penser; et que des habiletés spécifiques de la pensée critique se relient à des dispositions spécifiques de la pensée critique. Si ces suppositions sont vraies, elles suggèrent qu'un programme scolaire centré sur des habiletés mènerait les étudiant(e)s à vouloir et à pouvoir bien penser. On présente une définition de la pensée critique fondée sur un consensus parmi des experts qui font de la recherche; avance que les dispositions humaines sont ni cachées, ni inconnaissables ; décrit un procédé scientifique pour développer des moyens pour mesurer des habiletés cognitives et des dispositions humaines ; résume des résultats des recherches empiriques récentes qui explorent le rapport possible entre les habiletés de la pensée critique et la motivation ou la disposition à employer ces habiletés. Des études empriques indiquent qu'à toute fin printique les corrélations admises comme hypothèses ne sont pas évidentes. Il semblerait qu'un enseignement efficace inclurait des stratégies qui forment le caractère intellectuel et qui ne concentrent pas exclusivement sur l'amélioration des habiletés cognitives.

Keywords: critical thinking, disposition, assessment, test development, CCTST, CCTDI

"We have warp drive capability. Yes.
But where can warp drive take us?"
Anji, *Star Trek Insurrection*

1. Two Questions

The general consensus is that critical thinking (CT) *per se* is judging in a reflective way what to do or what to believe. The cognitive skills of analysis, interpretation,

inference, explanation, evaluation, and of monitoring and correcting one's own reasoning are at the heart of critical thinking (APA, 1990). Through practice, and with guidance from a good instructor, we can develop our thinking skills (like our artistic, athletic, or leadership skills) to the extent our natural abilities allow. But we should take care not to confuse the component skills with the activity itself. Critical thinking (CT) is judgment, reflective and purposive.

For decades college text books and K-12 educational policy emphasized learning skills (Marshall and Tucker, 1992).US government policy, which first recognized college level CT in its 1994 articulation of national educational goals, used the language of skills and abilities in that legislation (US Congress, 1994). So much attention has been paid to the skills used in CT that we risk trivializing CT, characterizing it as merely an assortment of techniques rather than as a complex, thoughtful, purposeful process of forming judgments using reasons and evidence (Paul, 1990).

What might be described as an over-emphasis on skills has been countered recently by a rebirth of interest in the dispositional side of thinking (Siegel, 1988; Paul, 1990; Esterle, 1993; Facione & Facione, 1992; Tishman and Andrade 1996, Ennis, 1996). The empirical identification of two factors as being involved when persons engage in tests of their CT tends to support the philosophical distinction between the skill dimension and the disposition dimension of the use of CT (Taube,1997). However, some theorists, like Paul, Tavris and Wade, include the disposition to use CT skills as part of their definition of CT, (Paul, 1990; Esterle, 1993). Carole Wade explains saying, "Carol Tavris and I . . . wanted to get in the willingness as well as the ability because a person can master CT skills without being the least bit disposed to use them." (Esterle,1993). By analogy to the properties of physical objects, some have argued that just like a cup that breaks must have been breakable, the fact that a person uses a skill is evidence that the person is disposed to use that skill, (Ennis, 1994; Perkins, 1993). Perkins and Tishman include the ability to exercise a given thinking skill as part of the meaning of being disposed to use that thinking skill (Perkins, 1993).

Hence the obvious first question, one with implications for teaching as well as for theory: Does demonstrable overall skill at CT correlate positively with the overall disposition toward CT? A second question goes further: Do any specific CT skills correlate with specific CT dispositions? Since more or less incompatible conceptualizations, as well as listings of CT dispositions, exist in the literature (see Baron, 1987; Lipman, 1987; Facione and Facione, 1992; Perkins, 1993; Ennis 1996), to clarify these questions let us first consider some commonplace examples of human dispositions, other than CT.

2. The Nature of Human Dispositions: Experiential Insights

Common experience shows that some people are courageous, others cowardly. Some are tenacious, others give up too easily. Some are trustworthy, others are unreliable. Some are more or less compassionate than others. At times we name

both the positive human disposition and its antithesis; at times we speak as if a human disposition might be stronger in one person than in another.

We think of courage, compassion, trustworthiness and the like as part of a person's character. Describing someone in terms of their dispositions expresses the person's habitual ways of acting. For example, an expert soccer player might be offensive-minded (prone to attack with the intent of scoring goals) or defensive-minded (prone to fall back with the intent of preventing the opponents from scoring) in her style of play. One parent might be permissive and another authoritative in their approach to discipline. A taxi driver might be talkative or sullen. An administrator, entrepreneurial or risk-averse. These are things people can know about other people.They are things we can say about ourselves. Knowing a person's dispositions allows us to predict, more or less, how the person is most likely to act or react in a wide variety of circumstances.

Consider compassion. In saying someone is compassionate, we mean that the person is sensitive to the needs of others, open to being touched by their suffering, and moved to act with the intent of ameliorating that suffering. There is a certain sensitivity, vulnerability, and responsiveness wrapped up in our everyday understanding of human compassion. This example does not imply that every human disposition has three elements; rather it shows that key components, if there are any, of a human disposition may, at times, be open to analysis and understanding. But not every compassionate person is a Mother Teresa. Truly saying that one person is more or less compassionate (reliable, collaborative, aggressive, studious, self-serving, communicative, easy-going, violent, decisive, playful, deceitful, etc.) than another is revealing. It shows that in everyday life, including in serious contexts like employment, family relations, and the courts, we have the ability to discern, comprehend, and compare human dispositional strengths and weaknesses.

Common human experience supports a theory of human dispositions that would characterize them as knowable tendencies, readily accessible to description, evaluation, and comparison by oneself and others. Namable and describable as a nexus of attitudes, intentions, values, and beliefs, one's dispositions are among the distinguishing features of one's character or personality. As such, human dispositions are not mysterious, unaccessible, hidden qualities.

John Dewey described the dispositional aspects of thinking as "personal attributes" (Dewey, 1933). In social psychology a disposition, when the word is used at all, is generally conceived of as an attitude or attitudinal tendency (Nunnally, 1978; Cronbach, 1990). In the physical sciences, where the word is almost never used, a disposition is a molecular property that objects manifest under certain conditions. For example, at a certain air pressure and temperature, hydrogen is disposed to boil. Ennis takes the physical science approach as paradigmatic (Ennis, 1996). Our research team takes the social science approach, with a nod toward Dewey, as our starting point. We propose to use the word 'dispositions' as ap-

plied to humans to refer to characterological attributes of individuals. As such, a human disposition is a person's consistent internal motivation to act toward, or to respond to, persons, events, or circumstances in habitual, and yet potentially malleable, ways.

3. The Correlation of Willing with Able: Experiential Insights

From weeding the garden to filing tax returns, life abounds with examples where skill exists but not the consistent internal motivation to use those skills. From writing bestsellers to being an effective leader, there are things some of us are strongly disposed toward doing but lack the requisite skills to accomplish. But perhaps CT is a special case. After all, when we consider our friends, co-workers, and family members, how many people do we know who are skilled at CT but apathetic or disposed not to approach problem solving by thinking? And how many people do we know who sincerely value and are well disposed to use CT, but simply do not possess the strong cognitive skills?

Experienced professionals in accounting, counseling, education, engineering, government service, health care, law, and management probably would not find such cases difficult to imagine. Unfortunately, there may be far too many examples of people who have the talent for skillful thinking but lack the consistent internal motivation to use that talent. And, like the weak student, far too many profess a disposition toward thinking but, regretfully, are not capable of generating much mental horsepower.

To the open-minded, imaginative, and intellectually adventurous, reflection on human experience can reveal the a rich array of possibilities. Yet, while we might wish to base our theories only on what we discern through everyday living, that basis, as the history of science and the history of philosophy show, can result in less than adequate understandings and misguided programs. To more fully explore reality one must move beyond comfortable descriptions of common experience and attempt to test, to the extent possible, one's hypotheses about the world in well-executed empirical studies.

4. The Empirical Research on the Concept of CT

Efforts to define and measure CT intensified throughout the last quarter of the Twentieth Century (Kurfiss, 1988; Norris & Ennis, 1989; Jones, 1993). In 1990, under the sponsorship of the Committee on Pre-College Philosophy of the American Philosophical Association, a cross-disciplinary international panel of 46 experts completed a two-year, multi-round, strict-method Delphi research project. which yielded a robust conceptualization of CT for purposes of instruction and educational assessment.

> We understand critical thinking to be purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as

> explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgment is based . . . . CT is essential as a tool of inquiry. As such, CT is a liberating force in education and a powerful resource in one's personal and civic life . . . . While not synonymous with good thinking, CT is a pervasive and self-rectifying human phenomenon. (APA, 1990)

Thus conceived, CT was characterized as a self-adjusting process of judging what to believe or what to do in a given context. In so doing a person engaged in CT uses a core set of cognitive skills to form that judgment and to monitor and improve the quality of that judgment (APA, 1990). The successful application of these core CT skills requires that one take into reasoned consideration the evidence, methods, contexts, theories, and criteria which, in effect, define specific disciplines, fields, and areas of human concern.

For many years CT testing focused primarily, if not exclusively, on measures of skills. To give only one paradigmatic example, *The California Critical Thinking Skills Test*, introduced in 1990, provides scores on sub-scales named *analysis*, *inference*, *evaluation*, *deductive reasoning*, and *inductive reasoning* (Facione, P., *et al.* 1990, 1998).

Many participants in the Delphi Project and others as well maintain that focusing only on CT skills is not adequate for instructional purposes. John Chaffee says, "A critical thinker is not only capable of reflecting, exploring, and analyzing but chooses to think in these advanced, sophisticated ways" (Esterle, 1993). Carol Tavris says, "To crystallize this spirit of thinking, to develop an understanding that there are other was of thinking about problems and other ways of thinking about our own behavior, including our 'feelings.' The question is: How do we get everybody to actually do it?" (Esterle, 1993).

While many endorse the importance of the critical spirit or the overall disposition toward CT, few take an empirical approach to measuring that disposition or exploring the elements of that overall disposition to value and utilize CT (Ennis, 1994; Facione, 1994; Salomon, 1994; Tishman, 1994, Ennis, 1996). My research colleagues and I decided to explore whether the disposition toward CT could be measured and analyzed empirically. In our research we characterize the overall *disposition* toward CT as the *consistent internal* motivation to engage problems and make decisions by using CT (Facione, Facione, & Giancarlo, 1996). In this we followed the 1990 Delphi research, which determined that while CT per se is defined as a form of cognition (judgment), to instill in students the disposition toward CT required an appreciation of the characterological profile one was trying to nurture. For this purpose the Delphi investigation offered a rich description of the ideal critical thinker.

> The ideal critical thinker is habitually inquisitive, well-informed, trustful of reason, open-minded, flexible, fair-minded in evaluation, honest in facing personal biases, prudent in making judgments, willing to reconsider, clear about issues, orderly in complex matters, diligent in seeking relevant infor-

> mation, reasonable in the selection of criteria, focused in inquiry, and persistent in seeking results which are as precise as the subject and the circumstances of inquiry permit. (APA, 1990)

Our research team's empirical refinement of the Delphi concept of the ideal critical thinker using the methods of experimental psychology, which are described below, began with the development of the *California Critical Thinking Disposition Inventory* (CCTDI) (Facione & Facione, 1992).

Unexpected support for the validity of the Delphi concept of CT, which grounds the CCTDI and the CCTST, came in the form of independent research sponsored by the United States Department of Education. Seeking to understand the Goals 2000 legislation, the Department of Education initiated a national survey of hundreds of educators, employers, and policy-makers to determine their priorities with regard to the communication and thinking skills expected of all college graduates (Jones, *et al.* 1994). The National Center for Post-secondary Teaching, Learning, and Assessment used the APA Delphi Research listing of CT skills and the seven dispositional factors identified in the development of the CCTDI as the expressions of CT for purposes of the national survey. The results, published in 1995, found strong national consensus that CT dispositions (and skills), characterized generally as expressed in the APA 1990 Delphi Report, should be essential outcomes of a college education (Jones, *et al.* 1995).

The 1990 Delphi description of the ideal critical thinker was further tested by Giancarlo (1996) using the California Q-sort method (Block, 1961). Twenty nationally recognized experts in adult CT individually sorted the 100 Q-sort items so that the result would characterize, in their opinions, the ideal critical thinker. The twenty were merged to form a prototype personality profile of the ideal critical thinker ($r=.80$, $N=20$, $p<.001$). [This shorthand means that the correlation "r" which can range from -1.00 to +1.00 was a very strong .80; the number of people sampled "N" was 20; and the probability of this result having occurred merely by chance "p" was less than 1 time out of 1000.] This prototype correlated with undergraduate students = CCTDI scores ($r=.36$, $N=91$, $p<.001$) and with CCTDI and Q-sort data collected about those students from their peers ($r=.32$, $N=91$, $p<.001$). These findings further confirm the strength of the 1990 Delphi consensus description of the ideal critical thinker and they support the use of CCTDI as a valid measurement tool relative to that definition.

Before we can cite empirical evidence in response to the question about the possible correlation of CT skills and the disposition toward CT, we must first be sure that we have valid and reliable measures of both. Section 5 below describes the process of tool development. Section 6 identifies the specific CT skills and the elements of the disposition toward CT that are described by the two measures used to gather the data cited in this paper.

5. Tool Development, Validity and Reliability

The process for developing valid and reliable tools to measure cognitive skills and dispositions in given population has been an established part of social and behavioral science research methodology since the middle of the Twentieth Century (Nunnally, 1978; Cronbach, 1990; Miller, 1991; Rosenthal and Rosnow, 1991). Decades of research undergirds the now familiar empirical paper and pencil methods used to gather data about a person's cognitive aptitudes and skills (e.g. multiple choice tests), and measures of a person's values, beliefs, intentions, opinions, attitudes, or inclinations (using the Likert style "agree-disagree" prompt). While there are potentially other valid and reliable ways to gather good educational assessment data, the use of a well-developed paper and pencil measure remains one fundamentally sound and sensible approach.

No approach to educational outcomes assessment is immune from the problems of establishing validity and reliability. Beyond the rigors of tool development and the challenges of executing a well-conceived assessment design, there are other problems: bad data resulting from the use of poorly developed tools or poorly executed educational assessment programs, honest concerns about the possible culture-bias or gender-bias of some instruments, appreciation of the test-taking problems of students with learning disabilities, misuse of modest testing results to bolster proposals for substantial changes in educational or employment policies, the national politics around high stakes educational testing, etc. It is no wonder that talk of testing leads quickly to many worries, some of which are quite reasonable.

The process of developing a good educational assessment tool of any kind begins with the construct, or idea, that one seeks to measure. The construct-validity of the instrument depends on how well that idea has been articulated and how well the tool captures that idea.The Delphi Report provided a wonderful opportunity for tool development, since it expressed a consensus construct of CT. The conceptualization used to structure an instrument, thus, was much more than simply a single individual's notion of CT. Further, the Delphi Report went well beyond a brief definition, for it offered lists of core CT skills and sub-skills and a robust expression of the positive side of the dispositional aspect of CT in its description of the ideal critical thinking.

The basic idea behind the design of educational assessment tools is that of representative sampling.The tasks, items, questions, used are developed as representative of the learning one is seeking to measure. Regardless of how valid it might or might not appear to be, typically, no one item on a test stands alone. Rather, it is the overall collection of items that is conceived of as covering the various elements that are most central to that learning. Gradations in item difficulty allow comparisons between those who have achieved a greater mastery of the learning and those who have a lesser mastery. So, if everyone is getting an item right or if everyone is getting it wrong, the item is probably of little value, since it

is failing to differentiate and might as well be eliminated. Another indication that an item should not be used is that students with a poorer overall mastery of the material tend to do better on that item than do students with a stronger mastery.

With the idea of representative sampling in mind, the next step in tool development is to write candidate items. In the case of a skills test, like the CCTST, they should be questions that are intended to elicit the use CT skills by persons in the target population to arrive at the designated answer. The items on the CCTST emerged from teaching CT since 1967 to thousands of students, having developed and used for that purpose many hundreds of educational skills test items, and having discussed with students their understanding and interpretation of the items. In the case of personality measures, like the CCTDI, one approach is to solicit "agree-disagree" responses to statements that vocalize or express the personal attitudes, intentions, values, and beliefs that are either central to or the antithesis of the disposition toward CT (Nunnally, 1978; Cronbach, 1990; Rosenthal and Rosnow, 1991). In 1990, using the rich description in the Delphi Report, our research team began the development of the CCTDI with the articulation of 225 statements expressive of one's disposition toward or away from using CT.

Before presenting a measurement tool, candidate test items are validated using a qualitative method involving talk-aloud and conversational strategies with persons in the target population. Many candidate test items are discarded or their wording is modified in this process. Typically this is done through focus group conversations with special attention paid to the exact phrasing persons in the target population use to express the attitudes, values, intentions, and beliefs (Triandis, 1980, Ferketich *et al.* 1993). This strategy provides a second form of validation, for it establishes whether or not the persons in the target population understand the test items in the way intended.

The next step is to pilot test the instrument in the target population. The pilot version often contains more items than will eventually be retained. We used 150 at this stage in developing the CCTDI. Statistical analyses of the responses of a sufficiently large and representative sample of test-taker allows for the elimination of items that fail adequately to discriminate among test-takers, items where the responses are inversely correlated with the overall scores on the test, and, in the interest of brevity, items that added little or nothing by way of further refinement of overall scores. This step yields the final tool, which is both more efficient and more reliable. In its final form the CCTDI includes 75 items. No single item, no matter how valid it may or may not appear on its face, can be thought of as standing entirely on its own, for it is always possible that a given test taker interpreted a given item in a way other than the one the test-maker intended.

A test is *valid* if it measures the thing it is intended to measure. For example, a bathroom scale measures weight. A test is *reliable* if it consistently gives the same results. (Cronbach, 1990). For example, a bathroom scale that gives different weights each time you stand on it, is not a reliable measure. All tests admit of many

potential sources of measurement error. A statistic that expresses the overall reliability of the test is called the Cronbach-alpha. A test that attempts to measure many different things where it is reasonable to expect that some students might have mastered some parts of the material but not other parts, will have a lower reliability. A test aimed at a clear, singular construct such that mastering that material would necessitate mastering all of its parts, should be expected, if it is a good test, to have an Cronbach-alpha in the neighborhood .80. Tests with alphas below .50 are suspect. Consider the construct "CT." To the extent that CT is a list of desperate skills, where it is possible to be good at some (say interpretation and analysis) but not others (say evaluation or explanation), what sort of alpha should we expect of a test of CT skills? Ennis and Norris at one time suggested .67 (Norris, 1989). To the extent that the disposition toward CT is not a singular thing, but actually a long list of 7, 12, 72 or more different dispositions and correlative dispositions, such that a person might have some and not others, one might wonder what an acceptable alpha should be for a measure of large sets of CT dispositions.

But there is a different approach. One that begins with the overall idea, and goes to constituent elements a posteriori. Given a valid and reliable test of a given construct, one can statistically explore that construct to see if there are different variables or factors working within it. In other words, instead of offering *a priori*, speculative analyses and lists of putative candidates, one can use factor analytic techniques on test data to achieve an *a posteriori* analysis. Factor analysis is "a broad category of approaches to conceptualizing groupings (or clusterings) of variables and an even broader collection of mathematical procedures for determining which variables belong to which groups" (Nunnally, 1978, p. 327). Factor-analytic techniques can yield lists of test items that cluster together because they are the responses to those items are so highly correlated with one another. By examining the test items in a given cluster, one can discern the conceptual meaning of that factor. Factors can thus be described and, if desired, named.

Each sub-set of items that form the factors on a given test can be thought of as a sub-scale which measures one of the elements of the overall construct that the test as a whole is measuring. The Cronbach-alphas of each sub-scale can be calculated. Often items are eliminated from the overall test so that the alpha of a given sub-scale can be increased. Because the sub-scales contain relatively few items as compared to the overall test, this usually does no serious damage to the alpha of the overall test. In this way scores can be reported as an overall result and in terms of sub-scales (Nunnally, 1978; Cronbach, 1990). This is how the seven scales for the CCTDI emerged, (Facione, 1992; Facione, N., Facione, P. & Giancarlo, 1994).

The question of *validity* can be expressed this way, "To what extent does a given test measure the right thing, namely the construct it purports to measure?" Starting with a well-articulated, research-based definition of the construct is very

important. Having independent confirmation of the validity of that definition gives further confidence that one has embarked on the right path. Then, based on understanding and experience, carefully writing well-targeted candidate items advances the process. Validating those items qualitatively, by asking representative members of the test's target population of test-takers how they interpret or understand the candidate items further strengthens test validity. If these earlier steps have been well executed so that the tool is focused on the right target, pilot testing the instrument and eliminating individual items that negatively correlate with the overall instrument scores contributes to our confidence in the validity of the remaining items. All of these steps were used in the development of the CCTST and the CCTDI.

After a test is developed, other support for its validity can be derived by examining the correlation of results of that test with previously existing, valid measures of the same construct. High correlations of CCTST scores with the analytical and verbal sections of the Graduate Record Exam, for example, suggest that the two are measuring very similar constructs (See Table 1). As indicated, the Giancarlo Q-Sort research lends additional conceptual and experimental support to the validity of the CCTDI. Additional independent support for the two testing tools can be found in studies that compare them to other commercial tests (Pendarvis, 1996).

Table 1

CCTST Correlations with Other Measures

Variable	r	N	Group	p-value
GRE Total	.719	143	A	<.001
GRE Analytic	.708	143	A	<.001
GRE Verbsal	.716	143	A	<.001
GRE Quantitative	.582	143	A	<.001
ACT	.402	446	B	<.001
CCTDI Total	.41	193	D	<.05
CCTDI Total	.201	1557	B	<.001
Watson-Glaser CTA	.405	139	B	<.001
Watson-Glaser CTA	.544	65	C	<.001
SAT Verbal	.545	123	B	<.001
SAT Verbal	.55	333	E	<.001
SAT Math	.422	123	B	<.001
SAT Math	.44	333	E	<.001
College GPA	.20	473	E	<.001
Nelson-Denny	.49	42	E	<.001
Age	-.006	479	E	.449
Units Earned	.03	473	E	.262

A = Nursing Graduate Students – CCTST
B = Nursing Students – CCTST Entry
C = Nursing Students– CCTST Exit
D = English Speaking 10th Graders – CCTST
E = 1990 Validation Study – CCTST Pretest

Even if the tool one is using is both reliable and valid, the practical context within which social science research occurs presents several potential pitfalls to arriving at a valid and reliable repeatable result in any given study. Educational assessment of learning outcomes is a prime example. The context is replete with circumstances that can confound our measurement efforts. For example, one worry with low-stakes tests of cognitive skills are that test-takers may not be motivated to give their best efforts. For a variety of reasons students might guess at the items, perhaps arriving at the correct answers but not by thinking them through properly. One worry with "agree-disagree" response tests is that test-wise persons will discern the desired response and mark that choice even if it is not their honest response.

While worries, like social-desirability bias, cannot be overlooked, in behavioral science research these concerns are not necessarily insurmountable.Text books on research methods in the social sciences and educational testing discuss offer strategies for discovering whether a given potential threat is of genuine concern in a given case. They suggest ways to minimize, if not neutralize, those threats that are of concern. For example, one way to determine whether a given test is in fact being compromised because test-takers are simply responding with the social-desirable answer is to use the Crowne-Marlowe Social Desirability Scale (Crowne, 1960; Rosenthal, 1991). For example, ten items used on the CCTDI were tested with 902 women aged 18 through 90, and the responses to those items were found to have no relationship to social desirability (Facione, N. *et al.*,1998).

In social science testing, the context can make a major difference. The above evidence was gathered in the context of women responding to a health survey concerning their attitudes toward help-seeking for breast cancer symptoms. There is no evidence that the women did not answer truthfully. On the other hand, a Japanese translation of the CCTDI used with Japanese engineering and nursing students appears vulnerable to problems of social desirability. In Japanese culture *tateme* is responding to questions by saying what a person should do or should be. *Honne* is responding by saying what a person really wants to do or to be (July 1999, conversations with Professors Kiyoko Makimoko and Ariko Noji at UCSF.) Their translation of some CCTDI items into Japanese appears to elicits the *tateme* response.To improve the performance of the Japanese version of the CCTDI items must be phrased to evoke the *honne* response.

The most telling evidence that the CCTDI is not significantly subject to social-disability bias when used with North Americans is the surprisingly consistent and unflatteringly low score that emerge in response to the Truth-Seeking sub-scale. We describe truth-seeking as the courageous desire for best knowledge in any given situations, asking the tough questions, and being willing to follow reasons and evidence wherever they lead even if the result is contrary to one's own pre-conceptions and interests. The socially desirable responses to some of the items on that scale are so blatant that faculty frequently are amused by the item's candor. How could anyone fail to know that the desired response, particularly in the con-

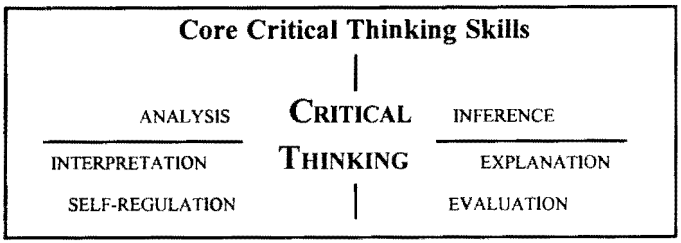
text of higher education, would be to disagree with statements like: "To get people to agree with me I would give any reason that worked." "Everyone always argues from their own self-interest, including me." "If there are four reasons in favor and one against, I'll go with the four." However, with a national mean on this scale falling solidly in the ambivalent range, it is clear from tens of thousands of cases that college students, whether or not they know the desired response, assert that they agree with these sorts of statements.

Frankly, it is not the validity of the test that worries those of us who have seen these data over the past seven years, rather it is the repeated evidence of ambivalence toward truth-seeking that most worries us.

6. Measuring Dispositions and Skills

Skills are manifest in performance. Persons with stronger skills tend to be able to perform a range of tasks requiring those skills with fewer mistakes. Lesser skilled persons make more errors on those same tasks. There are many ways to assess CT, including using performance appraisals, rating forms, rubrics, and portfolios (Facione & Facione, 1996a; Facione and Facione, 1996b).Research on psychological and educational testing indicates that well crafted multiple choice tests can validly and reliably measure higher order cognitive skills (Haldyna, 1994). The ball diagram with the names of the core CT skills is a useful reminder that we do not know the order in which one uses those skills in forming the judgment about what to believe or what to do. It also is a reminder that CT skills can apply not only to the question or evidence at hand, but to the products of the work of other CT skills. For example, we can interpret a set of data. We can analyze or evaluate that interpretation We can explain the analysis of that evaluation, or, using self-regulation, correct that analysis or that interpretation. We can revise the correction, interpret the analysis, etc. Furthermore, in this process we can revise our interpretation of the standards of evaluation, we can explain the problem in different ways, and we can re-evaluate our methodological assumptions. At times the complexities of good CT are evident when CT is carried on by groups, as when a management team is considering a business problem. Observation of the progression (and digressions) of the group's thinking, as manifest by what the participants say to one another during the deliberations, can show that sound, reflective judgments often come about in non-linear ways. Because there is so much we do not yet know about the cognitive complexity of CT in practice, instruction or assessment that emphasizes the exercise of a single CT skill in isolation from the possible influence of any of the others is problematic.

On the CCTST higher scores indicates greater skill at CT. Each CCTST item requires the test taker to form a judgment about the best response, from among those provided, to a question involving a more or less everyday situation or problem. No specialized knowledge is required, other than what can reasonably be expected of a person who has received a modest high school education and can



read at a 10th grade level. No technical vocabulary is used. To consistently answer correctly, the person must be skilled at inductive and deductive reasoning and at making correct analyses, inferences, and evaluations. The CCTST is a challenging test of CT, but it is not a test about CT (Facione *et al.*, 1990, 1998).Thus, a person's overall skill at CT can be arrayed along a continuum of test scores, which, in turn, permits statistical analyses, such as correlations. This provides a way to gather empirical data with regard to hypotheses such as CT skill is correlated with age, number of college credits earned, GPA, or scores on other educational or psychological tests. [Table 1 on p. 70 presents some of the correlations discovered to date.]

We characterized the disposition toward CT as the consistent internal motivation to use CT skills to decide what to believe and what to do. Some approach problems confident in their own ability to reason them through; others mistrust themselves as decision-makers, thinkers, or problem-solvers. Some people are open-minded about divergent ideas; others intolerant. Some are inclined to approach problems in diligent, focused, and systematic ways; others tend to be scattered, disorganized, and easily distracted. Some seek for evidence and reasons as they consider what to do; others eschew data and principled approaches, preferring rather to decide on the basis of impulse, whim, fashion, pressure, or caprice. Some seek answers as objectively as possible; others seem unable to step past personal bias, fear, self-interest, or preconception. Some see the complexity and subtlety of problems, noting multiple possible resolutions; others see things in stark, dualistic terms, as good or bad, right or wrong, true or false. Some people are curious as to the workings of things, wanting to know more about the problem than simply how to solve it; others are content just to know what to do, and prefer not deal with why or what if.

In general, people might be either positively or negatively disposed with regard to CT. And a third possibility is that they might not have formed a strong habit one way or the other; that is, they might be rather ambivalent in their tendencies or inclinations. Thus, one desirable feature of any measure of the overall CT disposition would be that it reflect the extent to which a person might be positive, ambivalent, or negative toward the use of CT.

Seven elements or aspects of the overall disposition toward CT emerged when statistical factor analytic techniques were applied in the initial development of the CCTDI. In their positive manifestations, these seven characterological attributes

were given the names *truth-seeking*, *open-mindedness*, *analyticity*, *systematicity*, *CT self-confidence*, *inquisitiveness*, and *maturity of judgment* (Facione & Facione, 1992). Seven intellectual bad habits can be thought of as the negative poles of each

The Disposition Toward Critical Thinking

Inquisitive
Systematic Judicious
Analytical Truthseeking
Open-minded Confident in Reasoning

intellectual virtue.The antithesis of the ideal would be a person who habitually approached problems, ideas, decisions, or issues being *intellectually dishonest* (e.g. in the use of data), *intolerant* (e.g. of opposing ideas), inattentive (e.g. to implications of proposals), *haphazard* (e.g. procedurally), *mistrustful of reason* (e.g. hostile toward sound scientific inquiry), *indifferent* (e.g. toward new findings), and *simplistic* (e.g. naively dualistic). In the final analysis, CT is about reflectively making sound judgments. So, while the intellectual virtues are obviously assets, the intellectual vices are even more obviously liabilities.The picture of the group talking with the names of the disposition sub-scales suggests that we can characterize the deliberations of groups of people, not just individuals, using these positive or negative attributions. Using 75 Likert style items to measure the seven characteristics and the overall disposition toward CT, the CCTDI gives us a means of gathering evidence that profiles the CT dispositions of individuals and groups.

7. The Hypotheses to be Tested

For purposes of empirical exploration, the two questions motivating this essay should be expressed as null hypotheses. Hypothesis #1: There is no correlation, positive or negative, between the persons' overall disposition toward CT and their strength or weakness in the CT skills. Hypothesis #2: No specific CT disposition is correlated, positively or negatively, with any specific CT skill. Using the CCTST to measure the skills and the CCTDI to measure the dispositions, the potential exists to gather some useful data that would tend to confirm or to disconfirm one or both of these null hypotheses.

8. Findings: Tenth Graders, Accounting and Nursing Professionals, and College Students

One early study provided evidence in support of a positive correlation between skills and dispositions at the macro level. It was a study of 10th grade high school

students in a school district in the Southwestern United States. In a sample of 193 students a positive correlation of r=.41 was found between CCTDI total scores and CCTST total scores (Giancarlo & Facione, N., 1994).This correlation was statistically significant at the p<.05 level, which tends to disconfirm the first null hypothesis and support the theory that there is a positive correlation between overall CT skill and CT disposition.

A statistically significant correlation, if squared, describes the percentage of the variance in one element of the correlation that might be attributable to the variance in the other element. Thus a .41 positive correlation between overall CT skill and the overall disposition toward CT suggests that up to 16.8% of the variance in CT skills test scores is potentially attributable to the differences in these students' CT disposition test scores. The same could be said the other way, that as much as 16.8% of the variance in CT disposition test scores is potentially attributable to the differences in these students' CT skills test scores. In other words, of all the many different things that could explain the students' performance on one of these two tests, up to 16.8% might, if we knew how, be explained by attribution to the element measured by the other of the two tests.

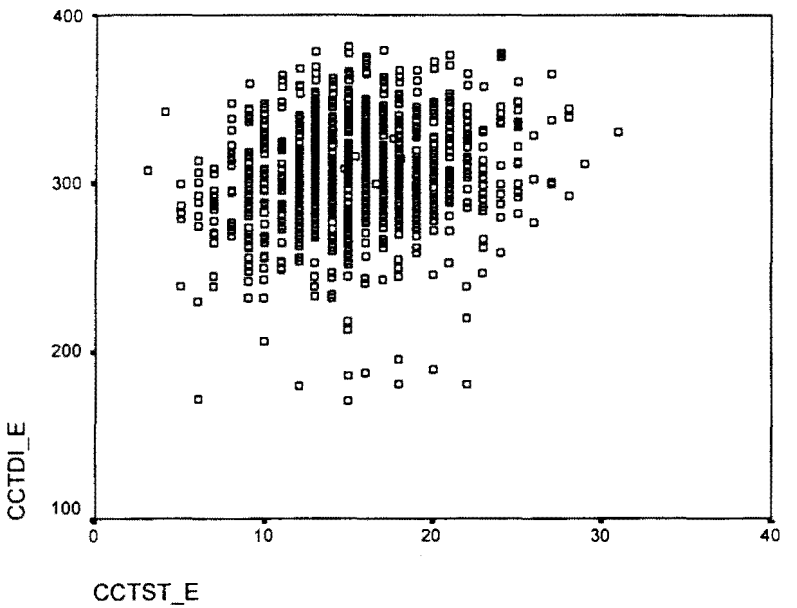
A 1998 unpublished analysis by statisticians at the California Academic Press of a dataset composed of 133 accounting professionals' scores showed an excellent mean CCTST score of 21.7 and a positive overall CCTDI means on all seven scales. Yet the correlation of overall skills test scores and overall disposition test scores as measured in this group of professionals with these instruments was only .091. Regardless of sample size, a number so close to zero would be interpreted as showing that there is virtually no correlation of skills and dispositions in this group, (Blohm, 1998). A study of 328 nursing undergraduates showed a statistically significant correlation (p <.01) of .318. This suggests that in this sample about 9% of the variance in skills test scores can be associated with (social scientists would say "explained by") the variance in overall CT dispositions (Colucciello, 1997).

By late 1997 sufficient data from multiple sites had been gathered to address both hypotheses in more than the preliminary ways mentioned above. Beginning in 1992 and on through the summer of 1997 data was collected by on-site investigators at 50 college level programs of nursing throughout the United States.The aggregate dataset, comprising 145 predominantly undergraduate samples, when completed, included information on 7,926 students. Students in this dataset ranged from entry level freshmen through graduate students, and included working adults returning to college, as well as traditionallyBaged undergraduates. To the best of this author's knowledge, this aggregate dataset is the largest collected to date within which it is possible to analyze potential relationships between CT skills, CT dispositions, and a variety of other academic, institutional, and demographic variables. And so, it was to this dataset that we turned in search of more substantial answers (Facione & Facione, 1997).

As with the earlier study of 10th graders, the national aggregate nursing student dataset leads us to reject the first null hypothesis. Analyses showed positive

correlations between the overall disposition toward CT and strength in CT skills. For example, 1557 nursing students completed both the CCTST and the CCTDI at entry to their college programs. The correlation of the total scores on the CCTST and the CCTDI in this 1557 person aggregation was r=.201, p<.001, (Facione & Facione, 1997). Figure 1 below is a graphic plotting of the correlation. Visually a strong positive correlation would be manifest more or less as a line running diagonally upward from the lower left corner to the upper right.The cloud-like splotch visible on Figure 1 helps illustrate how weak the correlation is in this large data set.

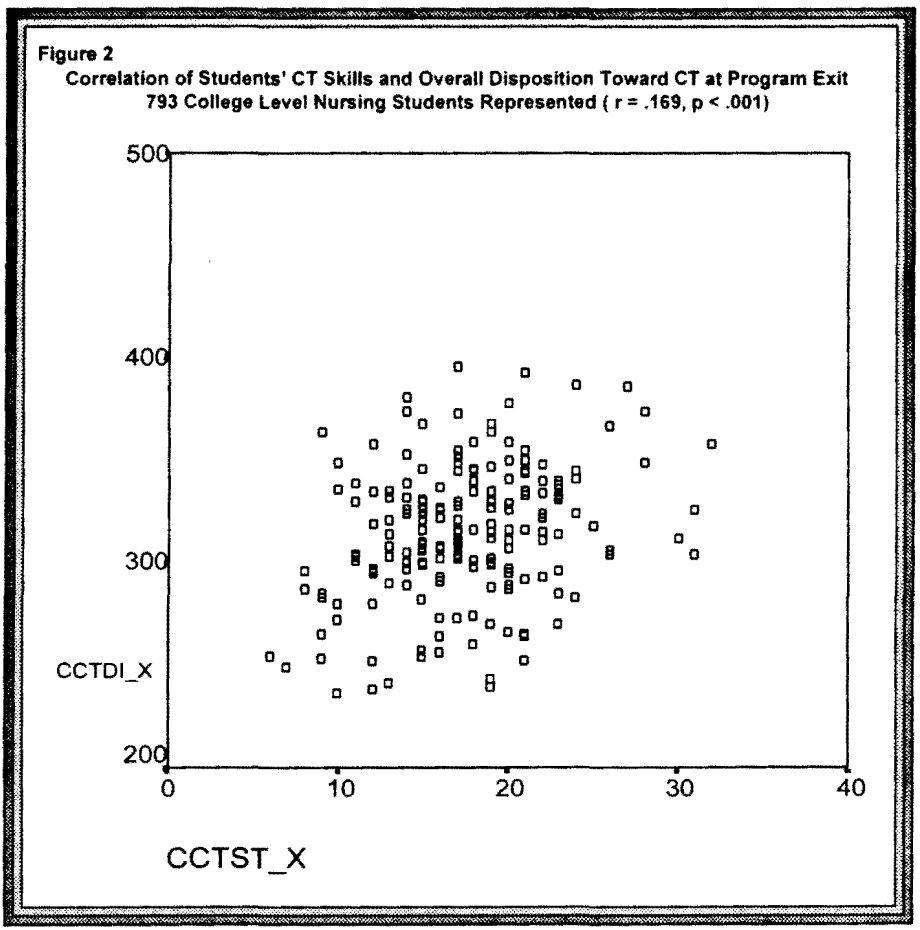
Figure 1
Correlation of Students' CT Skills and Overall Disposition Toward CT at Program Entry
1557 College Level Nursing Students Represented (r = .201, p < .001)



While statistically significant, we should be cautious not to over-interpret this finding. A correlation of r=.201 tells us that only about 4% of the variance in CT skills test scores potentially can be attributed to or associated with the variance of these college students' CT dispositions test scores, or visa versa. In other words, using these two measurement tools, 96% of the variance in CT skills in this sample of 1557 college students is not to be associated with or understood in terms of their CT disposition.

793 students, completed both the CT testing tools at exit from their nursing programs in the different colleges and universities.Again a positive correlation

(r=.169, p <.001) was evident (Facione & Facione, 1997). Figure 2 below plots this correlation. In Figure 2, the data points form a rather defuse cloud with only a hint of the diagonal. As with the entry level findings, however, the discovery of this statistical significant relationship should not lead us to believe that we have something of practical significance, for it fails to account for 97% of the variance. In other words, 97% of the difference between college students' CT disposition at program exit cannot be explained by reference to their CT skills at program exit, and *vice versa.*



The evidence from the entry and the exit data indicates even more clearly that the preponderance of the variance in CT skills is not potentially associated or attributable to variance in the overall disposition toward CT. The caveats for that generalization include: as CT is defined by the consensus construct developed in 1990, to the extent that the construct is measured validly and reliably by the two CT instruments used, insofar as the various studies aggregated into this analysis

were conducted appropriately, and to the extent that the college nursing student sample is representative of anything other than itself.

Interestingly, and perhaps as experienced instructors would expect, a somewhat stronger relationship was shown between the students' disposition score on the CCTDI at program entry and their skills test score on the CCTST at exit ($r=.233$, $p<.001$) (Facione & Facione, 1997). Those students with a stronger initial disposition toward CT showed greater development in CT skills by the end of their studies than did those with a weaker initial disposition toward CT.

The second question, whether any specific CT skill (e.g. analysis, or inductive reasoning) is correlated one-to-one with any specific CT dispositional factor (e.g. open-mindedness, truth-seeking, or systematicity), can be addressed by the data from the national aggregate sample. Depending on the possible pairings of skills and dispositional factors, there were between 1325 and 1428 students' cases to analyze statistically. With five CT skills and seven dispositional factors being tested, there are 35 possible pairings to consider. In thirty-three of the thirty-five possible relationships explored higher CT skills were generally correlated with stronger CT dispositions; but in all cases these correlations were very weak, never stronger than.194. Only the relationships between CT-self confidence and analysis that and between CT-self confidence and evaluation did not reach the threshold for statistical significance (Facione & Facione, 1997).

These findings tend to disconfirm the supposition that there is a one-to-one relationship between each specific CT skill and its supposed correlative disposition (as measured by the CCTST and the CCTDI). For if that were true, we would expect to find the skill of analysis, for example, to be strongly correlated to the disposition of analyticity, and not to the other six dispositional factors; and we would have expected to find analyticity to be strongly correlated only with skill in analysis, and not with the other skills.

In the light of these November 1997 empirical findings, it may be unwise to advance a theory that explicitly or implicitly pairs one and only one CT skill in a positive correlation with one and only one CT dispositional factor. While Perkins and Tishman did suggest such an approach, Ennis argued common sense grounds that it is both implausible and impractical (Perkins and Tishman, 1993; Ennis, 1996).The data presented here support Ennis' contention.

More importantly, what are we to make of the fact that skill at CT does not appear to be correlated with the disposition toward CT? What do findings like these imply for teaching CT and for the hope, widely shared among faculty, that disposition toward CT would assure that CT skills be used beyond the classroom?

9. Nurturing the Disposition and Building the Skill

Leading figures in the history of CT theory, persons who deserve the greatest credit for shaping the field, for bringing it international attention, and for leading

educational reforms toward a pedagogy based on teaching for and about thinking, have advanced theories that hypothesize a link between CT skills with CT dispositions. As was seen in the 1990 Delphi Research, some feel so strongly about the importance of CT dispositions that they want to include the disposition toward CT as part of the very meaning of 'critical thinking' (APA, 1990; Paul, 1990). The majority of 46 Delphi Research theoreticians, however, rejected this way of defining 'critical thinking.' The view that the majority espoused was more or less that the positive human characteristics of the ideal critical thinker are distinguishable from the critical thinking process itself. (APA, 1990). Many worried that a person might be skilled at CT but not fair-minded, and perhaps even unethical, in the use of those skills. Sensing in the impressive list of characteristics of the ideal critical thinker the potential for people to confuse "good CT" with "being a good (ethical) person," the majority of the participating theoreticians preferred a sharper boundary between ethics and CT. Some draw it very crisply precisely in order to distinguish between teaching skill at CT as contrasted with motivating students to use that skill in a consistent way (Fisher and Scriven, 1997).

In presenting a conceptualization of the ideal critical thinker, theoreticians are promoting a set of values relating to one's intellectual character, what might be called "intellectual virtues." In 1988 and 1989, the years when the Delphi research was being conducted, the participating theoreticians did not take the theory of CT dispositions much further, except that they did speculate that we would discover a strong positive correlation between CT skill and the disposition toward CT. Using the instruments currently available, that has not yet happened.

Skill and disposition are two separate things in people. Employers and educators prize both (Facione, Facione, & Giancarlo, 1996). A developmental perspective suggests that skills and dispositions are mutually reinforcing; and, hence, should be explicitly taught and modeled together (Kitchener & King, 1995). Common sense tells us that a strong overall disposition toward CT is integral to insuring the use of CT skills outside the narrow instructional setting. Motivational theory (Lewin, 1935) provides the theoretical grounds for the assumption that the disposition to value and utilize CT would impel an individual to achieve mastery over CT skills, being motivated to close the gap between what is valued and what is attained.

At the most practical level, learning follows motivation. We are best at learning what we most need and want to know. Thus, engendering the desire to use CT as a favored means of problem solving and decision making prepares the ground for teaching and learning the CT skills. Perhaps it is this linkage between motivation and learning that responds best to the insight that CT skill and the disposition to use CT should come together in practical and important ways. Speculation about a strong positive correlation does not capture the insight. But to assert that there is a pragmatic connection of tremendous value to learning and to living is not a philosophical mistake. In fact, the success of the entire enterprise of using reason,

rather than myths, magic, and mystery, is predicated on the pragmatic insight that we must nurture the dispositions even as we teach the skills.

Critical Thinking Pedagogy

Problem-Framing and Problem-Solving

- Model CT Skills and Dispositions
- Reward good CT
- Challenge poor CT
- Create a Climate of Reasoned Inquiry
- Diversify Judgment Contexts
- Engage Students in CT
- Reflect on the Use of CT

If the disposition toward CT can be described as an intellectual virtue, perhaps the analogy can be extended a bit and, without minimizing the centrality of developing CT students' CT skills, teaching CT could be thought of as including the building intellectual character. We worry that our educational programs at all levels, from K-12 through college and into adult and professional development, will fail if they focus only on skills to the neglect of the consistent internal motivation to use those skills in the appropriate circumstances. While necessary, it will not be sufficient only to strengthen students' skills on how to interpret, analyze, infer, explain, or evaluate. To teach for thinking, one must nurture truth-seeking, open-mindedness, analyticity, systematicity, intellectual curiosity, confidence in the proper use of reasons and evidence, and maturity of judgment.

> *If we were compelled to make a choice between these personal attributes and . . . some degree of technical skill in manipulating special logical processes, we should decide for the former.* (John Dewey, 1933)

Even though each person's CT ability has limits, we can still develop our skills to the extent that our ability permits. We can also expand the arena's within which we are adept at using our skills. CT is not limited in its content. Rather it is the process of judging, in a reflective way, what to do or what to believe. Thus its connections to professional judgement, and to thinking about problems and questions in science, ethics, and civic life are straightforward. Teaching for and about CT includes broadening the range of kinds of problems and decision making contexts within and about which we are willing and able to exercise our CT. In everyday situations and in every domain of knowledge or professional practice, good CT involves attending to the contexts, theories, methods, evidence, and standards within which problems are framed and decisions formed.

We would suggest that to be maximally effective teaching for and about CT should be aimed at expanding the opportunities as well as the content areas with

which to engage in informed, reflective problem framing and problem solving. One powerful tool for nurturing the disposition toward CT in students and co-workers is by modeling it. Stephen Brookfield says, "something . . . has been absent from a lot of [the] work on how you teach CT, and that is the crucial importance of educators modeling it in their own actions" (Esterle, 1993). If we are not truth-seeking, open-minded, and the rest, in our thinking with students, colleagues, and family about problems and decisions, is it not unreasonable of us to expect more of them?

10. Conclusion

We can trace western science and philosophy to the ancient Greeks who created a culture based on the commercial and civic utility of using reason to solve problems and make decisions. For them the intellectual virtues were as important as the civic and physical virtues. Although much of the educational and corporate rhetoric today is fueled by a pervasive and uncritical faith in high technology as an educational and economic panacea, wiser voices still whisper that information and skills alone cannot guarantee success in the workplace or in school. People must also be disposed to use what they have learned. *Educational and professional success require developing one's thinking skills and nurturing one's consistent internal motivation to use those skills.* To imagine a powerful positive automatic correlation between CT skills and CT dispositions actually undermines the task at hand. If we want our students to be both willing and able to engage in CT, and we do, then we have to include both in school and professional development curricula, in our instructional assignments, and in our educational outcomes assessments. Why? Because being skilled does not assure one is disposed to use CT. And, being disposed toward CT does not assure that one is skilled.

References

American Philosophical Association. (1990). *Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction.* ERIC document ED 315-423.

Baron, J. (1987). Evaluating thinking skills in the classroom. In Baron, J. and Sternberg, R., (eds.), *Teaching Thinking Skills: Theory and Practice.* 221-47. New York: W. H. Freeman and Company.

Block, J. (1961).*The Q-sort Method in Personality Assessment and Psychiatric Research.* Palo Alto, CA: Consulting Psychologists Press.

Blohm, S. (1998). Personal correspondence with the author, June.

Colucciello, M. Critical thinking skills and dispositions of Baccalaureate nursing students—A conceptual model for evaluation. *Journal of Professional Nursing*, **13** (4), 236-45.

Cronbach, L. (1990) *Essentials of Psychological Testing.* New York: Harper & Row.

Crowne, D. & Marlowe D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Counseling Psychology.* **24** (4), 349-54.

Dewey, J. (1933). *How We Think: A Restatement of the Relation of Reflective Thinking to the Educational Process.* Lexington, MA: Heath Publishing. Originally published in 1910.

Ennis, R. (1987). A taxonomy of critical thinking dispositions and abilities. In Baron, J. and Sternberg, R., (eds.), *Teaching Thinking Skills: Theory and Practice.* 9-26. New York: W. H. Freeman and Company.

Ennis, R. (1994). Assessing critical thinking dispositions: Theoretical considerations, (DRAFT). Presented to the American Educational Research Association Meetings, New Orleans, LA.

Ennis, R. (1996). Critical thinking dispositions: Their nature and assessability. *Informal Logic*, **18** (2&3), 165-82.

Esterle, J. & Clurman, D. (1993). Conversations with Critical Thinkers. San Francisco: The Whitman Institute.

Facione, N., Dodd, M., Miaskowski, C., Paul, S. & Giancarlo, C. (1998). How women decide to seek evaluation of breast symptoms. (Breast cancer research progress grant 1KB0045.) Paper presented at the International Conference on Cancer Nursing. Jerusalem, Israel. August.

Facione, N. & Facione, P. (1996a). Externalizing the critical thinking in knowledge development and clinical judgment. *Nursing Outlook*, **44**, 129-36.

Facione, N. & Facione, P. (1996b). Assessment design issues for evaluating critical thinking in nursing. *Holistic Nursing Practice*, **10**, 44-51.

Facione, N. & Facione, P. (1997). *Critical Thinking Assessment in Nursing Education Programs: An Aggregate Data Analysis.* Millbrae, CA: The California Academic Press.

Facione, N., Facione, P. & Giancarlo (Sánchez), C. (1994) Critical thinking disposition as a measure of competent clinical judgment: The development of the California Critical Thinking Disposition Inventory. *Journal of Nursing Education*, **33**, (8), 345-350.

Facione, P., *et al.* (1990, 1998). *The California Critical Thinking Skills Test (CCTST): Forms A and B; and the CCTST Test manual.* Millbrae, CA: California Academic Press.

Facione, P. (1992). Empirical methods of theory and tool development for the assessment of college level critical thinking and problem-solving. Paper prepared for the Second Department of Education, NCES Study Design Conference on Collegiate Skills Assessment, Washington D.C.

Facione, P. & Facione, N. (1992). The California Critical Thinking Dispositions Inventory (CCTDI); and the CCTDI Test manual. Millbrae, CA: California Academic Press.

Facione, P., Facione, N. & Giancarlo, C. (1996). The motivation to think in working and learning. In E. Jones (ed.), *Defining Expectations for Student Learning.* San Francisco, CA: Jossey-Bass Inc.

Facione, P., Giancarlo (Sánchez), C., & Facione, N. (1994). Are college students disposed to think? Paper presented at the Sixth International Conference on Thinking, Boston, MA. ERIC Doc: ED 368-311.

Facione, P., Giancarlo (Sánchez), C., Facione, N., & Gainen (Kurfiss), J. (1995). The disposition toward critical thinking, *Journal of General Education*, **44** (1), 1-25.

Ferketich, S., Phillips, L. & Verran, J. (1993). Development and administration of a survey instrument for cross cultural research. *Research in Nursing & Health*, **16**, 227-30.

Fisher, A. & Scriven, M. (1997). *Critical Thinking: Its Definition and Assessment*. Point Reyes, CA: Edgepress.

Giancarlo, C. (1996). The ideal critical thinker: development of an expert q-sort prototype. Presented at the American Psychological Association Meetings, Toronto, Canada.

Giancarlo, C. & Facione, N. (1994). *A Study of the Critical Thinking Disposition and Skill of Spanish and English Speaking Students at Camelback High School*. Millbrae, CA: The California Academic Press.

Haldyna, T. (1994). *Developing and Validating Multiple Choice Test Items*. Mahwah, NJ: Lawrence Earlbaum Publishers.

Jones, E. (1993). *Critical Thinking Literature Review*. The National Center for Post-secondary Teaching, Learning and Assessment. University Park, PA: Pennsylvania State University Press.

Jones, E., Hoffman, S., Moore, L., Ratcliff, G., Tibbetts, S. & Glick, B. (1994).*Essential Skills in Writing, Speech and Listening, and Critical Thinking for College Graduates: Perspectives of Faculty, Employers, and Policy Makers*. Project summary, U.S. Department of Education, OERI Contract No. R117G10037. National Center for Post-secondary Teaching, Learning and Assessment. University Park, PA: Pennsylvania State University Press.

Jones, E., Hoffman, S., Moore, L., Ratcliff, G., Tibbetts, S. & Click, B. (1995). *National Assessment of College Student Learning: Identifying the College Graduates' Essential Skills in Writing, Speech and Listening, and Critical Thinking*. National Center for Educational Statistics, U.S. Department of Education, Office of Educational Research and Improvement. Washington, DC: OERI publication NCES 93-001.

King, P. & Kitchener, K. (1995). *Developing Reflective Judgment*. San Francisco, CA: Jossey-Bass Publishers.

Kurfiss (Gainen), J. (1988). *Critical Thinking: Theory, Research, Practice, and Possibilities*. ASHE-ERIC Higher Education Report No. 2. Washington DC: ASHE.

Lewin, K. (1935). *A Dynamic Theory of Personality: Selected Papers*. Translated by D. K. Adams & K. E. Zener. New York: McGraw Hill.

Lipman, M. (1987). Some thoughts on the foundations of reflective education. In Baron, J. and Sternberg, R., (eds.), *Teaching Thinking Skills: Theory and Practice*, 152-61. New York: W. H. Freeman and Company.

Marshall, R. & Tucker, M. (1992). *Thinking for a Living: Education and the Wealth of Nations*. New York: Basic Books.

Miller, D. (1991) *Handbook of Research Design and Social Measurement*. Newbury Park, CA: SAGE Publications.

Norris, S. & Ennis, R. (1989). *Evaluating Critical Thinking*. Pacific Grove, CA: The Critical Thinking Press.

Nunnally, J. (1978). *Psychometric Theory*. New York: McGraw Hill.

Pendarvis, F. (1996). *Critical Thinking Assessment: Measuring a Moving Target.* Rock Hill, SC: The South Carolina Higher Education Assessment Network.

Paul, R. (1990). *Critical Thinking.* Rohnert Park, CA: Sonoma State University Center for Critical Thinking and Moral Critique.

Perkins, D., Jay, E. & Tishman, S. (1993). Beyond abilities: A dispositional theory of thinking. *Merrill-Palmer Quarterly,* **39** (1), 1-21.

Rosenthal, R. & Rosnow, R. (1991). *Essentials of Behavioral Research.* New York: McGraw Hill.

Salomon, G. (1994). To Be or Not To Be (Mindful)? Presented to the American Educational Research Association Meetings, New Orleans, LA.

Siegel, H. (1988). *Educating Reason.* New York, NY: Routledge.

Taube, K. (1997). Critical Thinking Ability and Disposition as Factors of Performance on a Written Critical Thinking Test. *Journal of General Education.*

Tishman, S. (1994). Thinking Dispositions and Intellectual Character. Presented to the American Educational Research Association Meetings, New Orleans, LA.

Tishman, S. & Andrade, A. (1996). Thinking dispositions: a review of current theories, practices, and issues. Cambridge, MA. Project Zero, Harvard University.

Triandis, H. (1980). Values, attitudes, and interpersonal behavior. In H. E. Howe (Ed.), *Beliefs, Attitudes, and Values: Nebraska Symposium on Motivation 1979,* 195-259. Lincoln: University of Nebraska Press.

United States Congress. (1994). *Goals 2000: National Goals for Education Act.* Washington, DC: U.S. Government Printing Office.

Peter A. Facione, Dean of the College of Arts and Sciences,
Dean of the Division of Counseling Psychology and Education
Santa Clara University, Ramos Center, Santa Clara, CA 95053 U.S.A.
pfacione@mailer.scu.edu