DOCUMENT RESUME

ED 454 248 TM 032 851

AUTHOR Edman, Laird R. O.; Bart, William M.; Robey, Jennifer; Silverman, Jenzi

TITLE The Minnesota Test of Critical Thinking: Development, Analysis, and Critical Issues.

PUB DATE 2000-08-00

NOTE 24p.; Paper presented at the Annual Meeting of the American Psychological Association (108th, Washington, DC, August 4-8, 2000).

PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.

DESCRIPTORS *Critical Thinking; Factor Analysis; *Factor Structure; Higher Education; Preservice Teacher Education; *Preservice Teachers; *Reliability; *Test Construction

ABSTRACT

The Minnesota Test of Critical Thinking (MTCT) has been designed to measure both critical thinking (CT) skills and a key disposition of critical reasoning: the willingness to evaluate arguments that are congruent with one's own goals and beliefs critically. The MTCT uses a taxonomy of CT skills derived from the American Philosophical Association's "Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction" (1990). Preservice teachers in training (n=210) were administered one of two forms of the MTCT by random assignment. Initial results indicate an overall Cronbach's alpha for form A of 0.76 and for form B of 0.69. These levels of internal consistency are perhaps appropriate in testing a construct that is itself multi-factor, and the levels are in the upper range when compared with other tests of CT. Examination of the correlation matrix of the subscales as well as the factor structure of the test indicates support for a hypothesized structure of CT with three aspects: metacognitive, analytic, and communicative. The instability of the subscale scores indicates the need for caution in interpretation, however. These results indicate the MTCT has potential for measuring CT skills, but could benefit from further revision and refinement. The results also indicate the need for increased research into the structure of CT. (Contains 29 references.) (Author/SLD)

The Minnesota Test of Critical Thinking: Development, analysis, and critical issues

Laird R. O. Edman, William M. Bart, Jennifer Robey, and Jenzi Silverman

The University of Minnesota

August, 2000

Please send requests for further information to William Bart, Educational Psychology, 211 Burton Hall, University of Minnesota, 178 Pillsbury Drive SE, Minneapolis, Minnesota 55455.

The Minnesota Test of Critical Thinking: Development, analysis, and critical issues

Laird R. O. Edman, William H. Bart, Jennifer Robey, and Jenzi Silverman

ABSTRACT

The Minnesota Test of Critical Thinking (MTCT) has been designed to measure both critical thinking (CT) skills and a key disposition of critical reasoning: the willingness to critically evaluate arguments which are congruent with one's own goals and beliefs. The MTCT uses a taxonomy of CT skills derived from the American Philosophical Association's Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction (1990). This taxonomy defines critical thinking as "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" (p. 3).

210 pre-service teachers-in-training were administered one of two forms of the MTCT by random assignment. Initial results indicate an overall Cronbach's alpha for form A of .76 and for form B of .69. These levels of internal consistency are perhaps appropriate in testing a construct which is itself multi-factor, and are in the upper range when compared with other tests of CT. Examination of the correlation matrix of the subscales as well as the factor structure of the test indicates support for a hypothesized structure of CT into three aspects: a metacognitive aspect, an analytic aspect, and a communicative aspect.

The instability of the subscale scores indicates the need for caution in interpretation, however. These results indicate the MTCT has potential as a valuable instrument in measuring CT skills, but could benefit from further revision and refinement. The results also indicate the need for increased research into the structure of CT.

Introduction

Over the past two decades, the focus of education has changed from curricular content to curricular outcomes, with a major emphasis on helping students learn to think critically (Edman, 1996; Fisher & Scriven, 1997; Klaczynski, Gordon, & Fauth, 1997; Halpern, 1998; Tucker, 1996). By 1995, most colleges and universities had included critical thinking (CT) skills as an important educational objective in their goal statements, and many accrediting agencies included measurable gains in critical thinking skills into their accreditation criteria (Facione & Facione, 1995).

This emphasis on teaching critical thinking necessarily leads to the need for reliable and valid ways of testing critical thinking. For example, the National League of Nursing has mandated all accredited nursing programs must teach CT to their nursing students and must empirically verify the efficacy of their CT instruction through testing (Rane-Szostak & Robertson, 1996). The assessment of CT is also at the heart of research on CT, for what cannot be measured cannot easily or convincingly be empirically studied. However, the measurement of CT is fraught with difficulty (Ennis, 1993; Tucker, 1996) and has proven to be one of the most difficult aspects of CT research.

Just as in the arena of intelligence testing where there is controversy over definitions, operationalizations, and thus over test construction, so also with CT testing. Because there is no standard definition of CT, the type of test one develops to test for CT depends heavily upon one's definition of the construct. If CT is defined as a set of reasoning competencies, then a measure of those competencies should suffice. However, most theorist and practitioners see CT as more than a set of reasoning competencies. The complex, probably multi-dimensional nature of CT makes simple tests of inductive and deductive logic unsatisfactory.

In order to inform pedagogy, research, and assessment, several CT theorists have proposed taxonomies of CT skills which elaborate the skills and aspects included in the term "critical thinking" (Dick, 1991; Ennis, 1987; Glaser, 1941; Paul, 1993). These taxonomies contain a great deal of overlap in their conceptual presentation of CT, but as of yet there has not been any empirical verification of the elements of CT. However, in 1990 the American Philosophical Association proposed a taxonomy of CT skills which was the result of a two-year Delphi study which included the input of 46 leading theorists and researchers in the field of CT pedagogy and assessment (American Philosophical Association, 1990). This panel defines CT as "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" (p. 3). The taxonomy of CT skills and subskills devised by this panel has the advantage of the combined expertise of the theorists on the panel, and as such is the most authoritative taxonomy of CT skills available.

The skills and subskills of CT, as delineated by the APA Delphi Study, are:

1. Interpretation
    - Categorization
    - Decoding Significance
    - Clarifying Meaning

2. Analysis
    - Examining Ideas
    - Identifying Arguments
    - Analyzing Arguments

3. Evaluation
    - Assessing Claims
    - Assessing Arguments

4. Inference
    - Querying Evidence
    - Conjecturing Alternatives
    - Drawing Conclusions

5. Explanation
    - Stating Results
    - Justifying Procedures
    - Presenting Arguments

6. Self-Regulation
    - Self-Examination
    - Self-Correction

There is widespread theoretical and empirical agreement, however, that critical thinking ability cannot be separated from a person's disposition to use that ability (Facione & Facione, 1995; Halpern, 1998; King & Kitchener, 1994; Klaczynski, Gordon, & Fauth, 1997; Paul, 1993; Perkins, Jay, & Tishman, 1994; Sa, W. C., West, R. F., & Stanovich, K. E., 1999). The relationship between thinking skills and the disposition or propensity to use them has been extensively examined, and several theorists posit that effective critical thinking is a function of two components: the competencies to perform specific cognitive operations, and the metacognitive skill and propensity to evaluate evidence independently of one's own goals and beliefs--to be open minded and objective (Kardash & Scholes, 1996; Klaczynski, Gordon, & Fauth, 1997; Stanovich & West, 1997). It is not enough for the critical thinker to have the skills to use reason when considering ill-defined problems. The critical thinker must also desire to use the skills even in situations in which reasonable reflection may lead to discomfort or difficult decisions on the part of the thinker. That is, the thinker must be willing to use critical thinking skills "against" even her or his own opinions and biases. This is what it means to be intellectually honest or to have intellectual integrity, oft-cited CT dispositional traits (Ennis, 1987; Facione, 1990; Paul, 1993).

If the disposition to use CT skills is an essential component of CT, a test of CT should incorporate assessing this dispositional element into its design. However, the currently available

standardized tests of CT measure the construct primarily as a set of reasoning skills divorced from the disposition to use the skills, and they have had only limited success in assessing CT. The current widely used tests of CT have been critiqued as having poor psychometric properties, of relying on limited conceptions of CT, of including confusing or ambiguous questions, and of lacking adequate empirically-based construct validity (Behrens, 1996; Fisher & Scriven, 1997; Follman, 1993; Harris & Clemmons, 1996; Jacobs, 1995; Moss & Koziol, 1991; Rane-Szostak & Robertson, 1996; Tucker, 1996). Many educators and theorists have called for new and better instruments for assessing CT ability (Ennis, 1993; Fisher & Scriven, 1997; Tucker, 1996).

The Minnesota Test of Critical Thinking (MTCT) has been designed to measure both CT skills and a key disposition of critical reasoning: the willingness to critically evaluate arguments that are congruent with one's own goals and beliefs. Using the taxonomy of CT skills listed in the APA Delphi study (APA, 1990), the MTCT is designed to employ an approach akin to Michael Scriven's multiple-ranking methodology, a methodology that creates dense response sets in a relatively limited amount of time (Fisher & Scriven, 1997). Using this methodology the authors hope to devise a test of CT that more fully and adequately assesses CT abilities and dispositions than do the currently available standardized tests of CT.

Methods

Participants

The participants in this study were 210 students from a wide spectrum of academic disciplines engaged in a post-baccalaureate teacher-training program at a large midwestern university. Their participation was voluntary and they received four extra credit points in the

educational psychology course in which they were enrolled for completing the test. No demographic information was gathered on the participants.

Instrument

A publication of the American Philosophical Association entitled "Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction" (1990) served as a guide in the development of this test. This publication is the end result of a series of interactions by a panel of experts using the Delphi Method. This method is an iterative process by which a group of experts responds to a series of questions in a thoughtful manner with the ultimate goal being a consensus opinion on an issue of some weight. This particular effort was aimed at creating "a consensus on the role of critical thinking (CT) in educational assessment and instruction" (p. 1, 1990). The final list of critical thinking skills includes: (a) interpretation, (b) analysis, (c) evaluation, (d) inference, (e) explanation, and (f) self-regulation. These six skills constitute the basis of the scales of the instrument used in this study.

Following the analysis of critical thinking assessment by Fisher and Scriven (1997), scenarios were created that address issues or controversies and that provided the basis for the assessment of the six critical thinking skills. The scenarios were intended to spark participant interests and to stimulate their own opinions on the issues at hand. The scenarios included issues of particular interest to educators and controversies of a more general nature.

The items address the six skills defined by the Delphi study. Each item was written in the form of a statement. Each statement contained a reference to an argument or element that might be considered important when making a judgment about the relative merits of a particular position on some issue. The participants were asked to rate each statement in terms of its importance to them in making such a judgment. Ratings were made on a 5-point scale: 1 = Not

at all important (NI), 2 = Somewhat important (SI), 3 = Important (I), 4 = Very important (VI), and 5 = Extremely important (EI). The following is an example item intended to assess the Interpretation subscale and that was used in conjunction with a scenario about the practice of retaining students who do not meet some criteria for passing on to the next grade:

To determine what the principal and your advisor mean by "in the best interest of the students"	NI	SI	I	VI	EI

Scoring the items required the development of a key. For each item, at least two of the developers of the test decided on an anchor point for the item, NI, I or EI. A third rater resolved any discrepancies that occurred. Subtracting the anchor point for a given item from the subject's response to the item and squaring that difference produced the subject's score for that item. Subscale and total scores were subsequently calculated by summing up these item scores and dividing by the number of items used to make up the scale or the entire test. Thus, each person's score on the items could range from 0, if there was no discrepancy between the individual's responses and the scoring key, to 16, if the anchor point was NI or EI and the subject responded at the opposite end of the rating scale. . In the example above, EI or Extremely Important, was designated as the anchor point. An individual could have received a score of 0, 1, 4, 9, or 16, depending on whether his or her response was EI, VI, I, SI, or NI, respectively. Thus, lower scores indicate greater agreement with the scoring key, and higher levels of critical thinking.

Procedure

In this study, the two forms of the test were given at random to students who had agreed to participate in the study. The tests were interleaved and then handed out in class and the participants were told they could complete the tests on their own time. A page at the end of the test asked for the participant's name and for an estimate of how long it took to complete the test.

This sheet was torn off and handed in to the course instructor to enable the awarding of extra credit points.

Results

Reliabilities can be found in Table 1. The overall Cronbach's alpha for Form A was $\underline{\alpha}$ = .7640 and the alpha for Form B was $\underline{\alpha}$ = .6932. Split-half reliability for Form A was $\underline{\alpha}$ = .6224 for the first thirty-two items and $\underline{\alpha}$ = .7673 for the remaining 32 items. The first 31 items on Form B produced a split-half reliability of $\underline{\alpha}$ = .5738 and the remaining 30 items had an alpha of $\underline{\alpha}$ = .7292. The subscales on Form A produced Cronbach's alphas ranging from $\underline{\alpha}$ = -.2771 to $\underline{\alpha}$ = .5960. The reliabilities for the subscales on Form B ranged from $\underline{\alpha}$ = .1097 to $\underline{\alpha}$ = .6161.

Table 1

Reliabilities				
	Form			
	A		**B**	
Total	α	# of items	α	# of items
Cronbach's α	.7640	64	.6932	61
Split-Half	.6224	32	.5738	31
	.7673	32	.7292	30
Scales				
Interpretation	-.2771	7	.2061	6
Analysis	.5659	12	.3411	13
Evaluation	.4854	16	.6161	14
Inference	.5960	16	.4674	16
Explanation	.3122	2	.1236	4
Self-Regulation	.3913	9	.1097	8

Descriptive statistics can be found in Table 2. Independent samples t-tests on the subscale scores on Forms A and B produced only one significant difference. The mean score for Evaluation was higher on Form A than on Form B ($t = 6.92$, $df = 208$, $p = 000$).

Table 2

Descriptive Statistics									
	Form								
	A n = 104			B n = 106			Independent Samples t-tests		
Scales	# of items	M	SD	# of items	M	SD	diff	t	p
Interpretation	7	2.57	1.57	6	2.29	1.31	0.28	1.42	.16
Analysis	12	1.72	0.88	13	1.83	0.80	-0.12	-1.01	.31
Evaluation	16	2.26	0.92	14	1.44	0.79	0.82	6.92	.00
Inference	16	2.19	1.19	16	2.43	1.12	-0.24	-1.49	.14
Explanation	2	1.77	1.22	4	1.91	1.15	-0.13	-0.81	.42
Self-Regulation	9	2.23	1.30	8	2.08	1.03	0.15	0.95	.34
Total	64	2.11	0.69	61	1.97	0.58	0.14	1.64	.10

Tables 3 and 4 contain data regarding the correlations between the individual items that contribute to each subscale and the total scores for those scales. The median r, the r-range and the number of significant correlations between the items and the subscale score are given. On Form A, all of the subscales scores correlate significantly with a majority of the items that contributed to making up each scale, with all of the items correlating significantly with the Analysis and Explanation subscales. The subscale scores on Form B also correlate significantly with a majority of the items used to make up the scales, with all the items correlating significantly with the scale scores for Analysis, Evaluation, and Explanation subscales.

Table 3

Correlations between Items and Subscale Scores					
	Form A				
Scales	Median r	r-range	# of significant correlations	# of items in the scale	% of significant correlations
Interpretation	.28	-.35 - .74	5	7	71.4%
Analysis	.44	.30 - .56	12	12	100%
Evaluation	.34	-.19 - .65	14	16	87.5%
Inference	.35	.13 - .67	11	16	68.8%
Explanation	.77	.73 - .81	2	2	100&
Self-Regulation	.35	.03 - .72	7	9	.77.8%

Table 4

Correlations between Items and Subscale Scores					
	Form B				
Scales	Median r	r-range	# of significant correlations	# of items in the scale	% of significant correlations
Interpretation	.38	-.12 - .73	4	6	66.7%
Analysis	.38	.19 - .48	13	13	100%
Evaluation	.43	.14 - .64	14	14	100%
Inference	.36	.07 - .49	14	16	87.5%
Explanation	.44	.42 - .74	4	4	100%
Self-Regulation	.34	.17 - .72	6	8	75.0%

The next two tables, Table 5 and Table 6, show the correlations among the subscales on the two different forms of the test. All significant correlations on Form A are positive. The Analysis subscale is significantly correlated with all the other subscales on Form A. In addition, the Interpretation subscale is significantly correlated with the Explanation and Self-Regulation subscales and the Evaluation subscale is significantly correlated with the Inference and Explanation subscales. The results for Form B are similar, but not identical, to the results for Form A. Again, all significant correlations are positive. The Interpretation subscale is significantly correlated to the Inference and Explanation subscales. Analysis is significantly correlated with the Evaluation, Inference, and Self-Regulation subscales. The Evaluation subscale is also significantly correlated with the Inference and Self-Regulation subscales and the Inference subscale is significantly correlated with the Self-Regulation subscale.

Table 5

Correlation Matrix – Form A Subscales						
	Interpretation	**Analysis**	**Evaluation**	**Inference**	**Explanation**	**Self-Regulation**
Interpretation	1.00					
Analysis	.35**	1.00				
Evaluation	-.02	.24**	1.00			
Inference	-.06	.23**	.59**	1.00		
Explanation	.24**	.35**	.18*	.14	1.00	
Self-Regulation	.35**	.40**	.15	.09	.23**	1.00

* = correlation is significant at the .05 level; ** = correlation is significant at the .01 level.

Table 6

Correlation Matrix – Form B Subscales						
	Interpretation	**Analysis**	**Evaluation**	**Inference**	**Explanation**	**Self-Regulation**
Interpretation	1.00					
Analysis	-.04	1.00				
Evaluation	-.14	.51**	1.00			
Inference	.24**	.19*	.25**	1.00		
Explanation	.18*	.05	.12	-.09	1.00	
Self-Regulation	-.12	.47**	.59**	.17*	.06	1.00

* = correlation is significant at the .05 level; ** = correlation is significant at the .01 level

Tables 7 and 8 show the component matrices for Forms A and B using the subscales as variables. On Form A, two factors emerged. The first factor includes the Analysis, Explanation, and Self-Regulation subscales while the second factor includes the Interpretation (although it is negatively correlated with the factor), Evaluation, and Inference subscales. Form B produced three factors. The first factor includes the Analysis, Evaluation, and Self-Regulation subscales. The second factor includes the Interpretation and Inference subscales. The third and final factor includes only the Explanation subscale.

Table 7

Component Matrix – Form A		
	Component	
Scales	**1**	**2**
Interpretation	.492	-.601
Analysis	.760	-.184
Evaluation	.580	.650
Inference	.527	.703
Explanation	.596	-.153
Self-Regulation	.627	-.351

Table 8

Component Matrix – Form B			
	Component		
Scales	**1**	**2**	**3**
Interpretation	-.115	.885	.012
Analysis	.775	-.004	.017
Evaluation	.854	-.066	.075
Inference	.390	.562	-.560
Explanation	.116	.364	.866
Self-Regulation	.816	-.123	.053

Conclusion

The reliability indices for the MTCT, forms A and B, are not high by the standards of many psychological tests. However, reported reliabilities of CT tests tend to be low (Ennis & Norris, 1990; Loo & Thorpe, 1999; Norris, 1995; Watson & Glaser, 1994), with Alphas ranging

from .37 to .87. The Cronbach's Alpha scores reported for the MTCT are in upper range of those reported for other tests of CT. One possible reason for the general tendency of lower reliability scores on tests of CT may rest with the nature of the construct. Reliability scores depend on the unidimensional nature of the construct being tested. If the construct is multi-dimensional, then reliability indices will tend to be low. CT, much like intelligence, is most probably a complex, multi-dimensional construct. Different items that measure some aspect of CT ability may not necessarily correlate highly with each other, even if they both are well-constructed, valid items.

One the other hand, one must not dismiss the obvious implication of low reliability scores: that the items in the test do not all measure the same construct, or do so in an unstable way. Further item-analysis is called for in order to explore the issue of reliability. This question is one that involves both issues of item development and construct dimensionality. Further research in this area may not only help refine the MTCT, but more importantly, shed light on the nature of the construct of critical thinking.

The reliability scores on the subtests are more troubling, revealing room for improvement in the subscale items. As is mentioned above, the negative values of the Alpha's for the Interpretation subscale on form A are worrisome, as are the low values for the Explanation subscale on both forms. However, possibly due to the fewer items involved in measuring each subscale, subscale values generally have lower reliability estimates on many psychological measures. The authors of the Watson-Glaser Critical Thinking Appraisal caution test users against interpreting individual subscale scores on the WGCTA because of the instability of such scores (Watson & Glaser, 1994). The MTCT appears to reflect this instability as well, something which may reflect not only on the instrument but on the nature of the subscales themselves.

On a more positive note, a high number of items have significant positive correlations

with the subscales in which those items are embedded (see table 3). This indicates a high level of item reliability within the subscales.

Perhaps the most intriguing finding from the above data concerns a hypothesis about the structure of CT, at least as measured by these subscales. As mentioned in the introduction, theorists have posited a variety of possible structures or taxonomies for CT, but as of yet there has been little to no empirical support for these theorized taxonomies. The present researchers have posited a possible clustering of the Delphi study subscales into three aspects of CT: a metacognitive aspect, an analytic aspect, and a communicative aspect. The correlation matrices in tables 4 and 5 offer tentative support to these clusters. Those subscale correlations that are significant on both forms of the MTCT appear to cluster into the three hypothesized CT aspects. Analysis, evaluation, and inference cluster together--an "analytic" aspect of CT. Self-regulation and analysis also appear to cluster together--a "metacognitive" aspect of CT. Finally, interpretation and explanation also appear to cluster together--a "communicative" aspect of CT.

These hypothesized aspects of CT are consistent with previous theoretical work in CT (Ennis, 1987; Halpern, 1998). Metacognitive skills (such as the ability to reflectively consider one's own thinking processes), analytic or reasoning skills (such as the ability to evaluate the need for and quality of evidence or the reasonableness of an argument), and certain communication skills (such as critical reading and listening) have all been strongly associated with the ability to think critically. At the core of this ability are the reasoning skills found in analysis, evaluation, and inference. When self-regulatory skills are added to these reasoning skills, we begin to approach what is often referred to as critical thinking, rather than simply good deductive and inductive logic. The relationship of interpretation and explanation to the core reasoning and reflective skills is less clear, although we believe these skills are also valuable

parts of CT. One can make a case for the theoretical importance of these aspects to overall CT, but the empirical support from this study for the relationship of this aspect with the others is not strong.

As mentioned above, however, the instability of the subscales on the MTCT (and on most CT tests) should lead to caution in interpretation. The initial support for these clusters of skills found here lends impetus for further research along these lines, and perhaps with further refinement, the subscales may lead to better analysis of the components of CT. Again, however, these results may indicate questions concerning the operationalization of the construct and the subscales, as well as psychometric issues with the instruments.

The instability of subscale scores may indicate not just problems with the test, but problems with the theory of the overall construct of CT as posited by the APA Delphi study. The particular sub-components of CT as envisioned in the Delphi report may not hold up under empirical scrutiny. On Form A, Principal components analysis of the results revealed a general factor that loaded on all subscales. This result is consistent with Norris (1995) who in a study that examined 15 different CT measures also posits a general critical thinking factor emerging from the results of confirmatory factor analysis. Such a general factor was less obvious from the analysis of Form B, although the first factor in form B does underscore our assertion of the central importance of analysis, evaluation, inference, and self-regulation in CT.

The subscales, however, are not recoverable from the principal components analysis, an expected result that also is consistent with other research on tests of CT instruments (Ennis & Norris, 1990; Loo & Thorpe, 1999; Norris, 1995; Tucker, 1996). The subscales are too highly interrelated to emerge from this analysis. What was puzzling, however, is the variety of the patterns that the analysis revealed. Interpretable second and third components did not seem to

emerge from the data. This supports the conclusions above that further research on the subscales is needed, and further refinement of the subscale items is called for.

Assessing critical thinking is an important, and in some cases, high stakes undertaking. As more and more secondary and post-secondary institutions look to teaching and testing the critical thinking skills of students, reliable and valid assessments must be designed. The MTCT is one part of a battery of tests being designed to provide such reliable and valid assessment. Testing critical thinking skills is a difficult task, however. Issues which need to be resolved include the extent to which critical thinking is a domain--specific or general competency, the extent to which critical thinking comprises discrete skills (such as identifying assumptions, evaluating credibility, deduction, induction, and metacognitive elements such as self-monitoring and self-awareness of cognitive strategies) which can be taught and tested individually or interdependent aspects of a complex concept that cannot be disassembled without altering its nature (Moss & Koziol, 1991), and the nature and importance of critical thinking dispositions to the teaching, assessment, and practice of critical thinking.

The MTCT may be a useful instrument for testing CT abilities and researching the questions raised above. However, the results of the current study indicate the MTCT, while showing promise, would benefit from further revision and refinement.

References

American Philosophical Association. (1990). Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction. The Delphi Report: Research findings and recommendations prepared for the committee on pre-college philosophy. (ERIC Document Reproduction Service No. ED 315-423).

Behrens, P. (1996). The Watson-Glaser Critical Thinking Appraisal and academic performance of diploma school students. Journal of Nursing Education, 35, 34-36.

Dick, R. D. (1991). An empirical taxonomy of critical thinking. Journal of Instructional Psychology, 18, 79-92.

Edman, L. (1996). Teaching teachers to teach thinking. The National Honors Report, 16, 8-12.

Ennis, R. H. (1987). A taxonomy of critical thinking dispositions and abilities. In Baron, J., & Sternberg, R. J., Eds. (1987). Teaching thinking skills: Theory and practice. New York: W.H. Freeman and Company.

Ennis, R. H. (1993). Critical thinking assessment. Theory into Practice, 32, 179-186.

Ennis, R. H., & Norris, S. P. (1990). Critical thinking assessment: Status, issues, needs.

In S. Legg & J. Algina (Eds.), Cognitive assessment of language and math outcomes (pp. 1-42). Norwood, NJ: Ablex.

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

Facione, P.A. & Facione, N. (1995). The disposition toward critical thinking. The Journal of General Education, 44, 25-50.

Fisher, A, & Scriven, M. (1997). Critical thinking: Its definition and assessment. Point Reyes, CA: Edgepress.

Follman, J. (1993). Critical thinking and verbal ability. ACEHI Journal, 19, 71-76.

Glaser, E.M. (1941). An experiment in the development of critical thinking. AMS reprint, New York: AMS Press.

Halpern, D. F. (1998). Teaching critical thinking for transfer across domains. American Psychologist, 53, 449-455.

Harris, J. C., & Clemmons, S. (1996). Utilization of standardized critical thinking tests with developmental freshman. Paper presented at the National Conference on Research in Developmental Education, October, 1996.

Jacobs, S. S. (1995). Technical characteristics and some correlates of the California Critical Thinking Skills Test Forms A and B. Research in Higher Education, 36 89-108.

Kardash, C. M., & Scholes, R. J. (1996). Effects of preexisting beliefs, epistemological beliefs, and need for cognition on interpretation of controversial issues. Journal of Educational Psychology, 88, 260-71.

Klaczynski, P., Gordon, D., & Fauth, J. (1997). Goal-oriented critical reasoning and individual differences in critical reasoning biases. Journal of Educational Psychology, 89, 470-485.

Loo, R., & Thorpe, K. (1999). A psychometric investigation of scores on the Watson-Glaser Critical Thinking Appraisal new form S. Educational and Psychological Measurement, 59, 995-1003.

Moss, P. A., & Koziol, S. M. (1991). Investigating the validity of a locally developed critical thinking test. Educational Measurement: Issues and Practices, 17-22.

Norris, S. P. (1995). Format effects on critical thinking test performance. The Alberta Journal of Educational Research, 41, 378-406.

Norris, S. P., & Ennis, R. H. (1989). Evaluating Critical Thinking. Pacific Grove, CA:

Critical Thinking Press and Software.

Paul, R. (1993). Critical thinking: What every person needs to know to survive in a rapidly changing world. 3rd ed. Santa Rosa, CA: Foundation for Critical Thinking.

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

Rane-Szostak, D. & Robertson, J.F. (1996). Issues in measuring critical thinking: meeting the challenge. Journal of Nursing Education, 35, 5-11.

Sa, W. C., Stanovich, K. E., & West, R. F. (1999). The domain specificity and generality of belief bias: Searching for a generalizable critical thinking skill. Journal of Educational Psychology, 91, 497-510.

Stanovich, K. E., & West, R. F. (1997). Reasoning independently of prior belief and individual differences in actively open-minded thinking. Journal of Educational Psychology, 89, 342-357.

Tucker, R. W. (1996). Less than critical thinking. Adult Assessment Forum Archives, 6, 1-6

Watson, G. B., & Glaser, E. M. (1994). Watson-Glaser critical thinking appraisal Form S

manual. San Antonio, TX: Harcourt Brace.

TM032851

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: The Minnesota Test of Critical Thinking: Development, analysis, and critical issues

Author(s): Edman, L.R., Bart, W.M., Robey, J., and Silverman, J.

Corporate Source:	Publication Date:
The University of Minnesota	August, 2000

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, ***Resources in Education*** (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 1	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2A	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2B
Level 1 ↑ [X]	Level 2A ↑ [ ]	Level 2B ↑ [ ]
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, → please

Signature:	Printed Name/Position/Title:	
Organization/Address: 2415 Tamarack Drive, Decorah, IA 52101	Telephone: 538-387-0068	FAX:
	E-Mail Address: edmon@oneota.net	Date: 4/13/01

APA 2000

(over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, *or*, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

University of North Carolina at Greensboro
ERIC/CASS
201 Ferguson Building
PO Box 26171
Greensboro, NC 27402-6171

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
4483-A Forbes Boulevard
Lanham, Maryland 20706

Telephone: 301-552-4200
Toll Free: 800-799-3742
FAX: 301-552-4700
e-mail: ericfac@inet.ed.gov
WWW: http://ericfac.piccard.csc.com

EFF-088 (Rev. 2/2000)