

Analyses of the Critical Thinking Assessment Test Results

J. Theodore Anagnoson
Professor of Political Science
November, 2000

Over the period from June, 1999 through June, 2000, Prof. Garry administered the Watson-Glaser test and the California Critical Thinking Skills Test (CCTST) to 28 different classes ranging from new first year students to seniors about to graduate. The classes broke down into five groups:

1. **Critical thinking classes.** These included five sections of Philosophy 160, two of Philosophy 250, two of Political Science 155 (Critical Analysis of Political Communications), and two of Speech 176 (Argumentation). Each of these classes counts for Critical Thinking in the Basic Subjects block of General Education. Each class was administered both a pre-test and a post-test, one on the first class day of the quarter and the other during the 10th week of the quarter or exam week. A total of 399 students took the pre-test; 339 took the post-test. With all of the normal changes in adds and drops, plus student absences, the results included pre-test and post-test scores, plus background information, for 282 students in all. About half of these students took the W/G; half, the CCTST.
2. **Control groups for the critical thinking classes.** These included two sections of Speech 150, where 54 students took the pre-test, 44, the post-test, and 42 both tests. These students were evenly split between the two exams.
3. **First time frosh in Introduction to Higher Education Classes.** These included one section of Arts and Letters 101, two sections of Engineering and Technology 100, and one section of Natural and Social Sciences 101. A total of 131 students took the pre-test, evenly split between the two tests, but since the students were not exposed to a critical thinking course, they did not take the post-test.
4. **Seniors.** Eight senior seminars took the two tests, with just under 40% taking the W/G exam. The seminars included one section each of Child Development 496, English 496, Nursing 497, Art 401, Management 497, and POLS 496, as well as two sections of Philosophy 496. There were 108 students in all.
5. **Juniors/Transfers.** Several sections of junior transfer students comprised the last group. This included one section each of Natural and Social Sciences 301, Engineering and Technology 154, and Nursing 348A and 349A. A total of 67 students were in this group, again composed roughly evenly of students who took each exam.

Thus, a total of 453 students took the pre-test; 383 students took the post-test, including both the CT group (#1 above) and the CT control group (#2 above). Another 306 did one exam.

Watson-Glaser Test Results:

We consider three aspects of the global results – first, the average score and how it rates by national standards; second, the percentage of the students who took the test who scored above the median national score, i.e., at or above the 50th percentile; and third, the percent of students who scored in the bottom decile. In other General Education assessment exams, the percent of students in the bottom decile has been anywhere from 25% to 50%, in part because of students who did not try on the particular test and in part because of low levels of student skills, including reading.

Norming group.

The Watson-Glaser test does not have a national sample against which students can be normed. Instead the test includes results from 17 groups of adults and students, ranging from “lower-level management applicants” to “construction management applicants,” “bank employees,” “ministry candidates,” “police officers,” and “miscellaneous occupations.” The scores for the 50th percentile vary from 31 to 34.5, with one group scoring at 26. The median of the medians for the lower nine scores is 31.5, and we have used that as our 50th percentile. The same procedure was used for the 10th percentile, with the range being from 19.5 to 28.5, with the mean at 25.5. The median of the nine lowest 10th percentiles was 25, and we have used that as our 10th percentile cutoff. For the “median score and percentile” below, the percentiles chosen were from the “Nurse Managers and Educators” group, which had its 10th percentile at 25 and its median at 31.5. While the results would have been different with other groups, they would not have varied by more than a few percentiles, less than 5-10.

Watson-Glaser Results:

Group	Median Score And Percentile	Percent above the 50th Percentile	Percent Below the 10th Percentile
Critical Thinking			
Students	21 (1 st)	6%	78%
Control Group	21 (1 st)	7%	77%
Seniors	24 (5 th)	13%	58%
First Year Students	20 (1 st)	4%	82%
Juniors/Transfers	24 (5 th)	13%	54%
All	22 (3 rd)	7%	74%

These scores seem low, even compared with the Academic Profile exams and the Collegiate Assessment of Academic Progress (CAAP) Writing Skills Test, both of which were administered by the Ad Hoc General Education Assessment Faculty Working Group during 1999-2000.

Here are the post-test scores for the critical thinking group and the control group at the end of the quarter, after the CT group has had its critical thinking course:

Group (Post-test scores only)	Median Score And Percentile	Percent above the 50th Percentile	Percent Below the 10th Percentile
Critical Thinking			
Students	21 (1 st)	6%	79%
Control Group	21.5 (1 st)	5%	82%

There is virtually no change.

The average Watson-Glaser score differed significantly by the grade received in the course:

Grade	Mean	N
D	19.25	12
C	20.29	38
C+	19.00	15
B-	20.11	19
B	21.42	33
B+	20.23	13
A-	23.72	18
A	25.92	48
Total	21.99	196

The F value for a one-way analysis of variance testing the hypothesis that all the grades are equal was 6.20, with a probability of arising by chance of less than 0.00005.

Demographic groups.

The scores vary predictably by the demographic variables collected at the same time as the test was administered:

Average Watson-Glaser Score

Group	Father		Mother		
	Average Score	N	Average Score	N	
Parent's Education					
BA+	23.3	32	22.4	17	
BA	23.1	27	23.3	44	
Some college	22.3	46	22.3	41	
High school graduate	21.6	59	22.0	59	
Less than a high school graduate	21.3	95	21.1	98	
Ethnic group					
Caucasian	28.2	21	One way analyses of variance on both father's and mother's level of education were insignificant. The one for ethnicity resulted in a F value of 6.69, with a probability of arising by chance less than 0.00005.		
Latino	20.8	132			
Asian	23.0	65			
African-American	21.4	21			
Age					
15-19	20.9	97	A one-way analysis of variance on the age categories resulted in an F value of 3.88, with a probability of arising by chance of 0.004.		
20-21	21.0	48			
22-27	22.0	71			
28-34	24.9	29			
35-65	25.4	17			
First language					
English	23.9	106	A one-way analysis of variance on the language categories resulted in an F value of 5.74, with a probability of arising by chance of 0.0008.		
Spanish	20.4	105			
Chinese	21.9	22			
Other	21.0	39			

The results above indicate that the W-G scores are related in predictable ways – older students do better, students whose parents, particularly the father, are better educated do better (and vice versa), students whose first language is English do a bit better. Caucasians and Asians do a bit

better than Latinos and African-Americans, although the Caucasian group contains only 21 individuals.

Multiple regression.

A multiple regression of the simplest model predicting the W-G score is contained in Table 1. The grade in the critical thinking course is closely related to the score received (the Pearson correlation between the two variables is 0.37). Age relates positively to the W-G score, with older students in general receiving slightly higher grades. White students on the average are four points higher than they would be otherwise, and those whose first language was English score 1.59 points higher, although the latter is not quite significant. A substantial portion of the variance in the W-G score is **not** explained by the model (over 75%), a result not unusual in predictions of this nature. A number of other variables were not significant when they were tested against those seen here, including the parent's education (either higher than average, or lower), and other ethnicities.

CCTST Test Results:

For the California Critical Thinking Skills Test (CCTST) we again consider the average score and how it compares with students at other institutions, the percentage of students who took the test who scored above the national median, i.e., at or above the 50th percentile, and the percent of students who scored in the bottom decile.

Norming group.

The CCTST literature provides three norming groups. The first is an analysis of 781 tests completed by a representative group of college students in 1989/90. They were mostly juniors from a "comprehensive, urban, state university." These students had not completed a course in critical thinking. The second group was composed of 153 Masters' level nursing students. The third group is composed of 224 cadets enrolled in a law enforcement academy. Since the Masters' students scored considerably better than the other two groups, and the cadets could be high school students, the first group, which is most likely from another CSU institution, is used for comparison. The cutoffs are 15 for the median and 10 for the 10th percentile.

CCTST Test Results:

Group	Median Score And Percentile	Percent above the 50th Percentile	Percent Below the 10th Percentile
Critical Thinking Students	11.0 (16 th)	19%	37%
Control Group	11.5 (21 st)	21%	36%
Seniors	14.0 (42 nd)	49%	12%
First Year Students	10.5 (10 th)	10%	40%
Juniors/Transfers	12.0 (25 th)	33%	16%
All	12.0 (25 th)	26%	30%

These are, in general, better results than the Watson-Glaser exam, particularly for the seniors. The CT students and control group are at the beginning of the quarter, before they have taken the course. At the end of the course, their scores are as follows:

Group (End of the quarter)	Median Score And Percentile	Percent above the 50th Percentile	Percent Below the 10th Percentile
Critical Thinking Students	12.0 (25 th)	30%	16%
Control Group	9.0 (6 th)	18%	64%

As can be seen, the CT group improved, with a one point increase in the median score to go from the 16th to 25th percentile, and an improvement in the percentage of students scoring above the 50th percentile from 19% to 30%. The percentage scoring below the 10th percentile fell from 37% to 16%. The corresponding scores for the control group worsened. These are better results than we obtained from the W-G test.

The average CCTST score also differed significantly by the grade received in the course, although not by as much as the Watson-Glaser test:

Grade	Mean	N
D	10.00	3
C-	11.00	5
C	9.64	44
C+	9.75	8
B-	13.67	6
B	11.26	34
B+	12.50	10
A-	12.07	15
A	12.79	39
Total	11.32	164

The F value for a one way analysis of variance testing the hypothesis that all the grades are equal was 2.62, with a probability of arising by chance of 0.01.

Demographic groups.

The results are again fairly predictable for the different demographic groups:

Average CCTST Score:

Group	Father		Mother	
	Average Score	N	Average Score	N
Parent's Education				
BA+	14.3	32	14.3	22
BA	13.5	52	13.5	50
Some college	13.9	56	13.0	49
High school graduate	11.8	67	12.7	70
Less than a high school graduate	11.6	110	11.5	132

Ethnic group

Caucasian	15.8	30	One way analyses of variance on both father's and mother's level of education were significant (the W-G ANOVAs were not). For father's level of education, the F value
Latino	11.5	158	
Asian	12.8	65	

African-American	12.0	29	was 3.95, $p < 0.004$; for mother's, the F value was 2.18, $p < 0.07$. For ethnic group, the F value was 4.17, $p < 0.0011$, also significant.
Age			
15-19	10.4	110	A one-way analysis of variance on the age categories resulted in an F value of 13.8, with a probability of arising by chance less than 0.00005.
20-21	11.2	50	
22-27	14.3	88	
28-34	16.6	36	
35-65	13.1	32	
First language			
English	13.2	143	A one-way analysis of variance on the language categories resulted in an F value of 3.71, with a probability of arising by chance less than 0.012.
Spanish	11.3	112	
Chinese	14.2	16	
Other	12.8	58	

The relationships here are stronger than they are for the W-G test, judging from the analyses of variance.

Multiple regression.

A multiple regression of the simplest model predicting the CCTST score is contained in Table 1. Like the regression predicting the W-G score, the grade in the critical thinking course is closely related to the score received (the Pearson correlation between the two variables is 0.30). Older students receive slightly higher grades, although the result here is not statistically significant. White students on the average are 3.38 points higher than they would be otherwise, and students whose father's educational level was below the high school level were 1.32 points below where they would have been otherwise. The first language English variable used in the W-G regression was not significant here. Again, over 80% of the variance in the dependent variable is not explained by the model.

A number of other variables were not significant when they were tested against those seen here, including the parent's education (higher than average), other ethnicities, and whether the student is currently in a CT class.

Table 1
Multiple Regressions of Test Scores on Demographic Variables

Model	1	2
Dependent Variable	Watson-Glaser Score	CCTST Score
N	133	129
Constant	12.24*** (6.08)	5.91** (3.13)
CT Grade	1.62** 3.17	1.19** (3.11)
Age	0.16* (2.19)	0.05 (0.75)
Caucasian	4.14* (2.28)	3.38** (2.63)
First Language English	1.59 (1.61)	
Father's Education Low (HS, <HS)		-1.32* (-2.04)
R-square	0.22	0.19
F-value	8.97	5.64
Significance	(0.0000)	(0.0001)

Note:

t-values are in parentheses below each coefficient.

* Significant at the 0.05 level.

** Significant at the 0.01 level.

*** Significant at the 0.001 level.

Pre-test and Post-test Results.

The core of the critical thinking assessment is a comparison of the pre-test and post-test results for the CCTST and Watson-Glaser tests for those students who took both tests. As one might expect, given that the students took the pre-test the first day of class and the post-test during the 10th week, a smaller number of students took both exams than took either exam.

- For the Watson-Glaser test, 173 took both the pre-test and the post-test.
- For the CCTST, a total of 151 took both exams.

Watson-Glaser Results		Pre-test			Post-test		
Group	N	Mean	SD	t-test of Difference	Mean	SD	t-test of Difference
Critical Thinking	152	22.03	5.545	1.23	21.671	5.861	-1.07
Control	21	20.429	5.996	(0.2211)	23.048	4.924	(0.2873)

For the pre-test, the t-value of the difference between the CT group and the control group is 1.23, with a 0.22 probability of arising by chance. We provisionally conclude that there is no difference between the CT group and the control group before the CT group took the course. For the post-test, again there is no significant difference between the two groups, although the CT group scored lower on the post-test than the control group. The CT group's post-test score also slightly declined compared with the pre-test, although again this difference was not statistically significant.

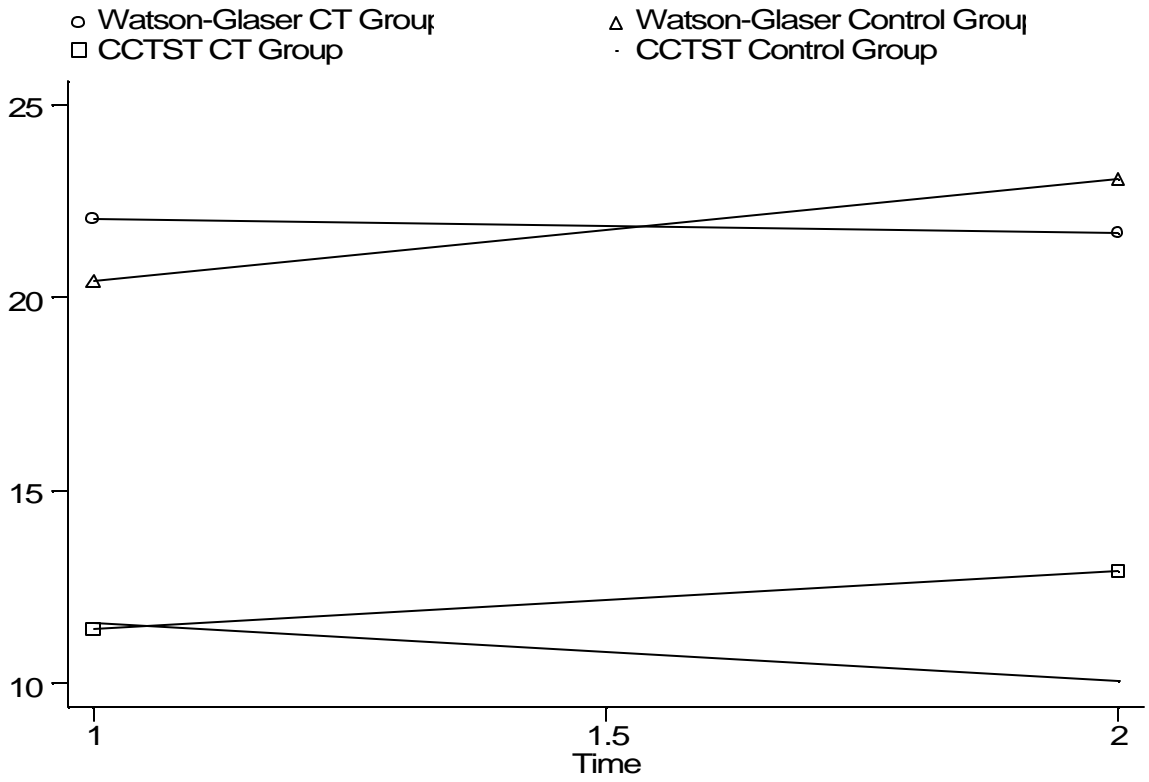
CCTST Results		Pre-test			Post-test		
Group	N	Mean	SD	t-test of Difference	Mean	SD	t-test of Difference
Critical Thinking	130	11.35	4.12	0.12	12.89	3.87	3.07
Control	21	11.24	3.36	(0.9030)	10.09	3.86	(0.0025)

For the pre-test, the scores are almost identical and the difference is not statistically significant. For the post-test, however, the critical thinking group has improved from 11.35 to 12.89, while the control group has fallen from 11.24 to 10.09, and the difference is highly statistically significant, with a t-value of 3.07 and a probability of arising by chance of less than 0.01.

A test of significance of the increase in scores for the CCTST from the pre-test to the post-test shows that the increase is highly significant, with a t-value of 3.48 and a probability of arising by chance of 0.0006, the lowest probability of any of the tests in this group.

The CCTST thus seems to be picking up on what is taught in CSLA CT courses and responds well to the difference between the treatment and control group.

Here is a graph showing the difference between pre-test and post-test for the Watson-Glaser and CCTST:



Watson-Glaser and CCTST Results