INSTITUTIONAL REPORT

CENTER FOR ASSESSMENT & IMPROVEMENT OF LEARNING

TENNESSEE TECH UNIVERSITY

©Tennessee Tech University, 2017

Contacts

Center for Assessment and Improvement of Learning

Payments & Contracts	Elizabeth Honeycutt	ehoneycutt@tntech.edu	931-372-3252
General Correspondence, Orders & Training	Dr. Elizabeth Lisic	elisic@tntech.edu or CAT@tntech.edu	931-372-3611
Reports/Analysis & Training	Dr. Kevin Harris	kharris@tntech.edu or CATreports@tntech.edu	931-372-3886
Project PI & Co-Director	Dr. Barry Stein	bstein@tntech.edu	931-372-3562
Project Co-PI & Co-Director	Dr. Ada Haynes	ahaynes@tntech.edu	931-372-3815

Center for Assessment and Improvement of Learning

244 Mathews Hall
80 West 8th Street
Tennessee Tech University
Cookeville, TN 38505

Phone: 931-372-3252
Fax: 931-372-3722
Email: CAT@tntech.edu

National Advisory Board	
Dr. John Bransford	University of Washington
Dr. Donald Deeds	Drury University
Dr. Peter Ewell	NCHEMS The National Center for Higher Education Management Systems
Dr. Michael Grant	University of Colorado
Dr. Gregory Light	Northwestern University
Dr. Patricia Turner	University of California

 ©Tennessee Tech University, 2017

The CAT Instrument

The CAT instrument is a unique tool designed to assess and promote the improvement of critical thinking and real-world problem solving skills. The instrument is the product of extensive development, testing, and refinement with a broad range of institutions, faculty, and students across the country. The National Science Foundation has provided support for many of these activities.

The CAT instrument is designed to assess a broad range of skills that faculty across the country feel are important components of critical thinking and real world problem solving. The test was designed to be interesting and engaging for students. All of the questions are derived from real world situations. Most of the questions require short answer essay responses and a detailed scoring guide helps ensure good scoring reliability.

The CAT instrument is scored by the institution's own faculty using the detailed scoring guide. Training is provided to prepare institutions for this activity. During the scoring process faculty are able to see their students' weaknesses and understand areas that need improvement. Faculty are encouraged to use the CAT instrument as a model for developing authentic assessments and learning activities in their own discipline that improve students' critical thinking and real-world problem solving skills. These features help close the loop in assessment and quality improvement.

Effectively Using the CAT Instrument

Assessment Models/Designs

The CAT instrument is adaptable to a variety of assessment goals and designs. We discuss these assessment goals and some of the more frequently used models below.

The CAT instrument can be used for a variety of assessment goals.

- Evaluate effects of college education
- Evaluate effects of a program of study
- Evaluate effects of a course
- Evaluate effects of informal learning experiences

There are a variety of assessment designs that can be employed with the CAT instrument. The CAT instrument is very adaptable to various research/assessment designs because the test is very sensitive to treatment effects and because the test can be used with all levels of college students without floor effects (students obtaining the minimum score possible) or ceiling effects (students obtaining the maximum score possible). These include:

- Pre-test/Post-test designs
    - Test students at the beginning and end of course or experience (with or without a control group).
    - Test students when they are freshmen and then again when they are seniors (true value added).
- Cross-sectional studies
    - Compare freshmen to seniors (typical value-added analysis).

 ©Tennessee Tech University, 2017

- Evaluate changes in program outcomes over time
    - Compare scores on the CAT after program improvements to established baseline scores that precede program changes.
    - Compare scores on the CAT to national norms over time and look for improvements.
- Evaluate changes in programs or courses by comparison to a control group.
    - Compare scores on the CAT for students who have had special courses/experiences to those for a control group who have not had the special courses/experiences.

Reducing Costs with Appropriate Sampling

We advocate a variety of practices to reduce the cost of testing without compromising the accuracy of the assessment. For example, various sampling strategies can be used to reduce the need to test all students. If that is not possible, then only a sample of the tests given might be scored. We discuss two accepted methods of sampling to ensure valid and representative results. However, we realize that the sampling techniques are not feasible at all institutions. Center staff will be happy to discuss these and other alternatives in more detail.

1. Random sampling: A subset of the student population of interest is randomly selected for testing/scoring. The larger the sample, the more confidence there is that the sample is representative of the population of interest. In a random sample, all students have an equal chance of being selected. This is not to be confused with a convenience sample that includes only those students who volunteer to take the test.
2. Stratified random sampling: The population is divided into subgroups (e.g., Arts & Sciences, Engineering, Education, etc.). A random sample of students within each subgroup is then selected. The number of students in each randomly sampled subgroup should be proportional to that group's proportion of the population. Stratification can help ensure a more representative sample with smaller sample sizes.

Sampling after Test Administration

In many institutions it is not possible to administer the test to a random sample of students within a class. In these situations, we recommend administering the test to the larger group and then randomly sampling tests from that group to score during the faculty scoring session. This procedure will allow institutions to achieve a more representative sample without greatly increasing the faculty time needed to score tests. We recommend having a minimum of 10 – 15 tests or pairs of tests per group (e.g., class, program of study, etc.).

Scoring Accuracy Checks

At various times during the year, we conduct analyses of scoring accuracy and provide feedback about the accuracy of scoring and, if necessary, specific recommendations for improving the accuracy of scoring on a question-by-question basis. These reports are sent separately from the institutional summary report.

 ©Tennessee Tech University, 2017

Example Assessment Designs to Use with the CAT Instrument

(These designs can easily be coded in the Local Code Field on the CAT Instrument)

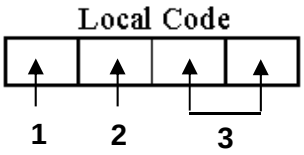
Objective	Model/Design	Sampling Procedure	Sampling Before Scoring	Advantages/ Disadvantages
Find courses or programs of study that Improve students' critical thinking	**Pre-test vs. Post-test** In selected courses or programs of study (matched students)	Administer to all students at the beginning and end of certain targeted courses or experiences	Randomly sample pairs of tests to score from each course or experience. (minimum of 10 matched pairs of tests per class)	A powerful and efficient design to evaluate specific courses and experiences (student IDs must match).
	(students not matched)		(min. of 15 pretests and 15 post-tests per class)	Less efficient & less powerful than above
	Pre-test vs. Post-test with Control Group In selected courses or programs of study (matched students)	Administer to all students at the beginning and end of certain targeted courses or experiences	Randomly sample pairs of tests to score from each course or experience. (minimum of 10 matched pairs of tests per class)	A powerful design to evaluate treatment effects relative to a control.
	(students not matched)		(min. of 15 pretests and 15 post-tests per class)	Less efficient & less powerful than above
	Treatment vs. Control	Administer to all students at the end of certain targeted courses or experiences	Randomly sample tests that will be scored after administering to a larger sample	Might be difficult to establish equivalence of treatment & control conditions.
How much is the institution or program of study improving students' critical thinking	**Freshmen vs. Upperclassmen** (value added) **Cross-sectional Study** (must equate groups)	Administer to a random sample of freshmen and seniors every year	Randomly sample tests that will be scored after administering to a larger sample	Might be difficult to establish equivalence of Freshmen and Upperclassmen if there is attrition.
Is the institution making progress in improving students' critical thinking	**Cross Sectional Study** of seniors over time (with or without national norm comparison)	Administer to a random sample of seniors (or all seniors) every year	Randomly sample tests that will be scored after administering to a larger sample	Would be necessary to establish the equivalence of samples over time.

©Tennessee Tech University, 2017

Using the Local Code Fields to Identify Assessment Design

A Local Code Field with 4 digits appears on the back of each test booklet. You should use this area to code subgroups in your population so that the data can be easily analyzed.

Recommended Use of Local Code



Column 1: Use to indicate the **Type of Design** for data included in the report.
Column 2: Use to indicate the student's specific **Treatment Condition** in the design.
Column 3: Use to indicate the particular class that was being tested **if a breakdown by class is desired**.

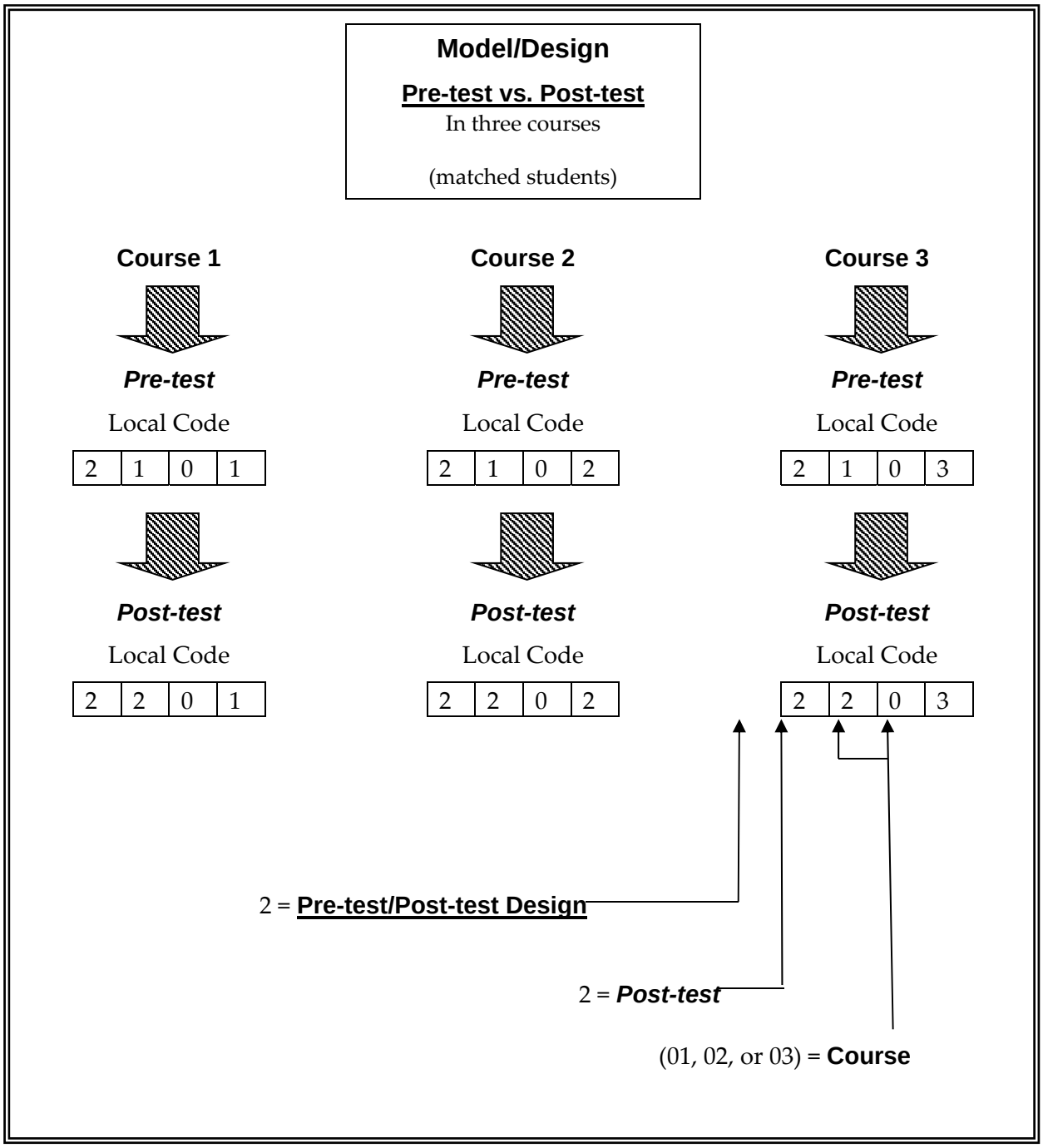
<table>
<tr><th>Column 1</th><th>Column 2</th><th>Column 3 & 4</th></tr>
<tr><td>0 (or blank) = No Breakdown</td><td>0 (or blank) = No Breakdown</td><td rowspan="7">0 (or blank) = No Breakdown

1 – 99 = code for each course or section tested if scores are to be compared by class</td></tr>
<tr><td>1 = Breakdown only by course</td><td>1 = Breakdown only by course</td></tr>
<tr><td>2 = Pretest/Post-test Design Matched Students
(identical ID #'s are used)</td><td>1 = Pretest
2 = Post-test</td></tr>
<tr><td>3 = Pretest/Post-test Design Not Matched Students</td><td>1 = Pretest
2 = Post-test</td></tr>
<tr><td>4 = Treatment vs. Control
(single post-test used)</td><td>0 = Control
1 = Treatment</td></tr>
<tr><td>5 = Lower Division vs. Upper Division Students</td><td>1 = Lower Division
2 = Upper Division</td></tr>
<tr><td>6 = Pretest/Post-test Design with Control Group Matched Students
(identical ID #'s are used)</td><td>0 = Control Pretest
1 = Control Post-test
2 = Treatment Pretest
3 = Treatment Post-test</td></tr>
<tr><td>7 = Pretest/Post-test Design with Control Group Not Matched Students</td><td>0 = Control Pretest
1 = Control Post-test
2 = Treatment Pretest
3 = Treatment Post-test</td></tr>
<tr><td>8 = Regression Design (use # of courses taken to predict CAT Score)</td><td colspan="2">These fields are left blank (contact us for more information).
Institution must supply a data file with student ID#s and number of courses completed.</td></tr>
</table>

©Tennessee Tech University, 2017

We can easily generate reports with breakdowns of data, if you use the coding scheme discussed on the previous page. We encourage you to contact us and discuss your plans for developing your local code before administering the test.

Example of How to Code Tests with Local Codes



 ©Tennessee Tech University, 2017

Correlations with Entering ACT and SAT Scores

Student scores on the CAT instrument correlate with their scores on college entrance exams like the ACT and SAT. These entrance scores can explain about 25% of the variability in student performance on the CAT instrument.

	ACT	SAT
CAT	0.501*	0.516*

* correlations significant, $p < .01$ (updated on 8/10/10)

We provide the following table to show how the average entering ACT/SAT score at an institution might impact upper division student performance on the CAT instrument at 4 year institutions.

CAT National User Norms
(Upper division undergraduate, 4 year institutions)

Average College Entrance Score*		Upper division
ACT (Composite)	**SAT (Verbal & Quantitative)**	**CAT Score (Estimated)**
13	620	10.79
14	680	11.93
15	740	13.07
16	780	13.83
17	830	14.78
18	870	15.54
19	910	16.30
20	950	17.06
21	990	17.82
22	1030	18.58
23	1070	19.34
24	1110	20.10
25	1140	20.67
26	1180	21.43
27	1220	22.19
28	1260	22.95
29	1300	23.71
30	1340	24.47
31	1380	25.23
32	1420	25.99
33	1470	26.94
34	1520	27.89

*Updated 8/10/10

©Tennessee Tech University, 2017

Your Institutional Report and Data File

CAT institutional reports provide information about your students' scores on the CAT instrument with descriptive information about sample demographics, mean score, minimum and maximum score, and standard deviation. In addition, the report provides a detailed frequencies breakdown of the distribution of answers (point values) for each question together with a general description of what the question is measuring. The mean score for each question and the percent of total points attained is also included. Current information about national norms is also provided. Additional comparisons are included as specified by the use of local codes.

A data file in Excel format is provided. This file contains the following information:

- Individual student responses for all questions on the demographics page and final scores for each test question.
- The file includes two spreadsheets, one sheet includes all student data, and the report sheet includes only students that did not have excessive missing data. The report is based on student data that is complete.

The institutional report and data file are uploaded to our secure FTP server to be accessed by the designated data owner at your institution. Contact Kevin Harris for more information (kharris@tntech.edu, 931-372-3886)

Variable	Type	Description
std_s1	Scale	Entrance Exam Score as entered by the institution
qpa	Scale	QPA as entered by the institution
testnum	Nominal	Test Booklet Number
stude1	Nominal	Student ID Number
loc-code	Nominal	Local Code as entered by institution
age	Nominal	Age
gender	Nominal	Gender (0=Male; 1=Female)
spanish	Nominal	Spanish/Hispanic/Latino (0=No; 1=Yes)
primary	Nominal	English is primary language (0=No; 1=Yes)
profi1	Nominal	Proficiency with English Language (1=Excellent; 2=Very Good; 3=Good; 4=Fair; 5=Poor)
standing	Nominal	Class Standing (1=Freshman; 2=Sophomore; 3= Junior; 4=Senior)
class	Nominal	University Standing (1=Undergraduate; 2=Graduate)
white	Nominal	Race: White (0=No; 1=Yes)
black	Nominal	Race: Black or African American (0=No; 1=Yes)
amer1	Nominal	Race: American Indian or Alaska Native (0=No; 1=Yes)
asian	Nominal	Race: Asian (0=No; 1=Yes)
nativ1	Nominal	Race: Native Hawaiian or Other Pacific Islander (0=No; 1=Yes)
other1	Nominal	Race: Other (0=No; 1=Yes)
q1f – q15f	Scale	Computed Score for each question.
total	Scale	CAT total score
q1 – q15	Scale	Computed Score for each question. (Rounded)
ptestnum *ploc-code*	*Nominal*	*Post test data for the corresponding variables listed above. (if applicable)*
pq1f – pq15f *ptotal* *pq1 – pq15*	*Scale*	*Post test scores for the corresponding variables listed above. (if applicable)*
report	Nominal	Case included in report (Y=Yes; N=No)

©Tennessee Tech University, 2017

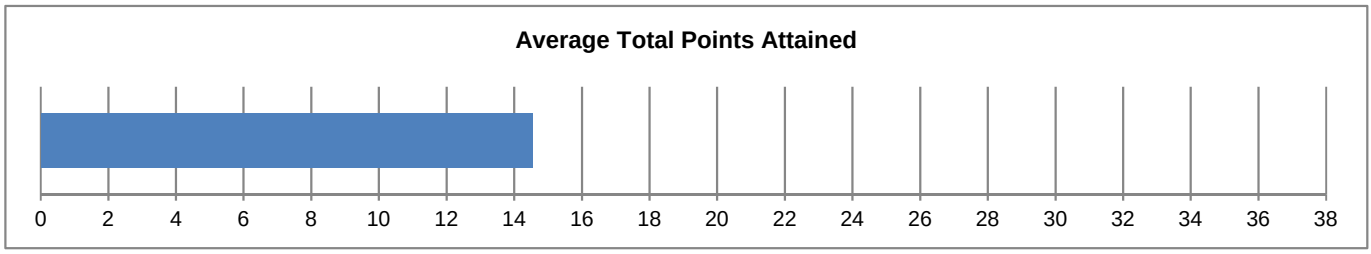
Sam Houston State University

CAT Institutional Report

August 2018 - All Students

CAT Overview: Descriptive Statistics for CAT Total Score
Sam Houston State University: August 2018 - All Students

	N	Min.	Max.	Mean	Std. Dev
CAT Total Score	513	1.00	28.00	14.54	5.17



CAT Demographics: Descriptive Statistics for Sample

		Freq.	Freq. %
Gender	**Male**	164	32.2%
	Female	345	67.8%
Class Standing	**Freshman**	5	1.0%
	Sophomore	25	4.9%
	Junior	206	40.5%
	Senior	273	53.6%
Class	**Undergraduate**	504	99.8%
	Graduate	1	0.2%
Age	**≤ 20 years**	126	26.0%
	21-25 years	311	64.3%
	≥ 26 years	47	9.7%

		Freq.	Freq. %
Race**	**White**	357	69.6%
	Black or African American	83	16.2%
	American Indian or Alaska Native	8	1.6%
	Asian	26	5.1%
	Native Hawaiian or Other Pacific Islander	6	1.2%
	Other Race	61	11.9%

**The cumulative percent may exceed 100% as students are allowed to select more than one category.

		Freq.	Freq. %
Proficiency with the English Language*	**Excellent**	384	75.0%
	Very Good	95	18.6%
	Good	30	5.9%
	Fair	2	0.4%
	Poor	1	0.2%

* Self-rated

	Freq.	Freq. %
Spanish/Hispanic/Latino Ethnicity	144	28.1%
Considered English primary language?	477	93.0%

CAT Breakdown: Frequency of Points Awarded for Each Question

Sam Houston State University: August 2018 - All Students

	Skill Assessed by CAT Question	Points Awarded	Freq.	Freq. %
Q1	Summarize the pattern of results in a graph without making inappropriate inferences.	**0**	206	40.2%
		1	307	59.8%
Q2	Evaluate how strongly correlational-type data supports a hypothesis.	**0**	212	41.3%
		1	197	38.4%
		2	70	13.6%
		3	34	6.6%
Q3	Provide alternative explanations for a pattern of results that has many possible causes.	**0**	268	52.2%
		1	139	27.1%
		2	75	14.6%
		3	31	6.0%
Q4	Identify additional information needed to evaluate a hypothesis.	**0**	259	50.5%
		1	140	27.3%
		2	69	13.5%
		3	37	7.2%
		4	8	1.6%
Q5	Evaluate whether spurious information strongly supports a hypothesis.	**0**	146	28.5%
		1	367	71.5%
Q6	Provide alternative explanations for spurious associations.	**0**	91	17.7%
		1	229	44.6%
		2	161	31.4%
		3	32	6.2%
Q7	Identify additional information needed to evaluate a hypothesis.	**0**	377	73.5%
		1	126	24.6%
		2	10	1.9%
Q8	Determine whether an invited inference is supported by specific information.	**0**	200	39.0%
		1	313	61.0%
Q9	Provide relevant alternative interpretations for a specific set of results.	**0**	242	47.2%
		1	213	41.5%
		2	58	11.3%
Q10	Separate relevant from irrelevant information when solving a real-world problem.	**0**	13	2.5%
		1	26	5.1%
		2	84	16.4%
		3	192	37.4%
		4	198	38.6%
Q11	Use and apply relevant information to evaluate a problem.	**0**	155	30.2%
		1	294	57.3%
		2	64	12.5%
Q12	Use basic mathematical skills to help solve a real-world problem.	**0**	131	25.5%
		1	382	74.5%
Q13	Identify suitable solutions for a real-world problem using relevant information.	**0**	234	45.6%
		1	185	36.1%
		2	59	11.5%
		3	35	6.8%
Q14	Identify and explain the best solution for a real-world problem using relevant information.	**0**	189	36.8%
		1	61	11.9%
		2	13	2.5%
		3	86	16.8%
		4	139	27.1%
		5	25	4.9%
Q15	Explain how changes in a real-world problem situation might affect the solution.	**0**	323	63.0%
		1	104	20.3%
		2	56	10.9%
		3	30	5.8%

Institutional/Departmental Profile

Sam Houston State University: August 2018 - All Students

Evaluate and Interpret Info	Problem Solving	Creative Thinking	Effective Comm.		Skill Assessed by CAT Question	Institution/Department	
						Mean	Avg. % of Attainable Points
X				Q1	Summarize the pattern of results in a graph without making inappropriate inferences.	0.60	60%
X			X	Q2	Evaluate how strongly correlational-type data supports a hypothesis.	0.86	29%
		X	X	Q3	Provide alternative explanations for a pattern of results that has many possible causes.	0.75	25%
	X	X	X	Q4	Identify additional information needed to evaluate a hypothesis.	0.82	21%
X				Q5	Evaluate whether spurious information strongly supports a hypothesis.	0.72	72%
		X	X	Q6	Provide alternative explanations for spurious associations.	1.26	42%
	X	X	X	Q7	Identify additional information needed to evaluate a hypothesis.	0.28	14%
X				Q8	Determine whether an invited inference is supported by specific information.	0.61	61%
		X	X	Q9	Provide relevant alternative interpretations for a specific set of results.	0.64	32%
X	X			Q10	Separate relevant from irrelevant information when solving a real-world problem.	3.04	76%
X	X		X	Q11	Use and apply relevant information to evaluate a problem.	0.82	41%
	X			Q12	Use basic mathematical skills to help solve a real-world problem.	0.74	74%
X	X			Q13	Identify suitable solutions for a real-world problem using relevant information.	0.80	27%
X	X		X	Q14	Identify and explain the best solution for a real-world problem using relevant information.	2.00	40%
	X	X	X	Q15	Explain how changes in a real-world problem situation might affect the solution.	0.60	20%
					CAT Total Score	14.54	38%

The map of skills covered by each question above is a suggested theoretical guide for interpreting results.

Upper Division CAT Means Comparison Report

Sam Houston State University: August 2018 - All Students

Evaluate and Interpret Info	Problem Solving	Creative Thinking	Effective Comm.		Skill Assessed by CAT Question	Institution	National		
						Mean	Mean	Probability of difference[a]	Effect Size[b]
X				Q1	Summarize the pattern of results in a graph without making inappropriate inferences.	0.60	0.67	***	-.15
X			X	Q2	Evaluate how strongly correlational-type data supports a hypothesis.	0.86	1.21	***	-.35
		X	X	Q3	Provide alternative explanations for a pattern of results that has many possible causes.	0.75	1.35	***	-.61
	X	X	X	Q4	Identify additional information needed to evaluate a hypothesis.	0.82	1.41	***	-.52
X				Q5	Evaluate whether spurious information strongly supports a hypothesis.	0.72	0.73		
		X	X	Q6	Provide alternative explanations for spurious associations.	1.26	1.56	***	-.36
	X	X	X	Q7	Identify additional information needed to evaluate a hypothesis.	0.28	0.82	***	-.90
X				Q8	Determine whether an invited inference is supported by specific information.	0.61	0.68	***	-.16
		X	X	Q9	Provide relevant alternative interpretations for a specific set of results.	0.64	0.93	***	-.41
X	X			Q10	Separate relevant from irrelevant information when solving a real-world problem.	3.04	3.14	*	-.10
X	X		X	Q11	Use and apply relevant information to evaluate a problem.	0.82	1.11	***	-.45
	X			Q12	Use basic mathematical skills to help solve a real-world problem.	0.74	0.82	***	-.18
X	X			Q13	Identify suitable solutions for a real-world problem using relevant information.	0.80	1.18	***	-.40
X	X		X	Q14	Identify and explain the best solution for a real-world problem using relevant information.	2.00	2.29	***	-.16
	X	X	X	Q15	Explain how changes in a real-world problem situation might affect the solution.	0.60	1.15	***	-.56
					CAT Total Score	14.54	19.04	***	-.80

[a]. * p<.05 **p<.01 ***p<.001 (2 –tailed) Does not Account for entering ACT/SAT.

[b]. Mean difference divided by pooled group standard deviation.
(0.1 - 0.3 = small effect; 0.3 - 0.5 = moderate effect; >0.5 = large effect)

The map of skills covered by each question above is a suggested theoretical guide for interpreting results.