Salisbury University CCTDI & CCTST Critical Thinking, Interdependence Among Disciplines, & Intellectual Curiosity Report, Fall 2015 & Spring 2016

This report, authored by SU office of University Analysis, Reporting & Assessment (UARA) staff, discusses Critical Thinking, Interdependence Among Disciplines, and Intellectual Curiosity-related survey data collected during both fall 2015 and spring 2016 GULL Week sessions.

Executive Summary

Background and Findings

1. Faculty and UARA staff agreed that California Critical Thinking Disposition Inventory (CCTDI) and California Critical Thinking Skills Test (CCTST) are aligned with General Education student learning goals 1.1. Critical Thinking, 2.2. Interdependence Among Disciplines, and 3.3. Intellectual Curiosity, as well as those goals' student learning outcomes.
2. The CCTDI instrument has 75 Likert-style items with an overall score as well as 7 scale scores. Similarly, the CCTST instrument has 34 multiple choice items with an overall score as well as 7 scale scores. Unfortunately, reliability and factor analyses of the fall 2015 CCTDI and spring 2016 CCTST SU student response data were not possible due to inability to access the raw item response data from the instruments' publishers.
3. The CCTDI and CCTST instruments were developed to be valid and reliable:
    a. CCTDI and CCTST validity:
        i. Content Validity: alignment with the American Philosophical Association (APA) Delphi Study's definition of critical thinking
        ii. Communication and Cultural Validity: development of forms of the instrument across languages based on feedback from researchers, professional translators, and native speakers
        iii. CCTST Item Analysis: item analysis has also been performed over the last 40 years to determine the item difficulty of items in the item pool
        iv. CCTST Scale Validity: scale analysis where item alignment to scales was supported by factor loadings from .300 - .770
        v. CCTST Criterion and Construct Validity: increases on scores on the test following critical thinking interventions; positive correlations with related measures of GRE Total Score: r = .719, p < .001; GRE Analytic r = .708, p < .001; GRE Verbal r = .716, p < .001; GRE Quantitative, r = .582, p < .001)
    b. CCTDI and CCTST reliability:
        i. Test-Retest reliability values of instruments meet or exceed .80
        ii. CCTDI instrument and scales demonstrate reliability (α = .71 - .80 in pilot study)
        iii. CCTST instrument and scales demonstrate reliability (KR-20 scores ≥ .70 in validation sample)
4. The results of our administration of the CCTDI and CCTST instruments supported its validity and reliability:
    a. Validity of scores:
        i. CCTDI Criterion and Construct Validity: SU student overall scores on this instrument had positive correlations with the SU students' related measures of SAT Verbal scores, r = .363 (p < .001); SAT Math scores, r = .169 (p < .05); and SAT total scores, r = .304 (p < .001)
        ii. CCTST Criterion and Construct Validity: SU student overall scores on this instrument had positive correlations with the related measures of SAT Verbal scores, r = .498, Math scores, r = .495, and total scores, r = .560 (all p < .001) as well as High School GPA, r = .178 and SU Cumulative GPA, r = .133 (all p < .05)

   b. CCTDI and CCTST reliability scores for this administration at SU in fall 2015 and spring 2016, respectively, were not evaluated due to additional cost for the analyses and lack of access to the raw data.
5. In general, the demographics of the students that took the CCTDI and CCTST instruments were similar to the non-test-takers, but due to the nature of the GULL Week sampling method there were groups that were not well represented.
6. In general, SU students' average overall score on the CCTDI instrument (295 ± 27 SD out of a range of 70 - 420 possible; n = 515) in fall 2015 aligns with "Positive" mindsets for critical thinking. Although the SU students' average scores indicated a "Positive" mindset for the aligned attribute for most of the scales, in both Truth-seeking and Systematicity scales there is evidence for a need for improvement since those scales' average scores indicated "Inconsistent/Ambivalent" mindsets.
7. In general, SU students' average overall score on the CCTST instrument (72.2 ± 6.0 SD out of a range of 50 – 100 possible; n = 831) in spring 2016 aligns with "Moderate" critical thinking skills. The average percentile of SU's students in this sample, when compared to the aggregate sample of CCTST Four Year College Students, is 35.4 (SD = 23.6). Although the SU students' average scores indicated "Moderate" critical thinking skills for the aligned scales - and therefore improvement is needed in all 7 scales - in particular, the scales Evaluation, Deduction, and Explanation should be prioritized for improvement.

Action Items

1. Evaluate the need to revise the related SU general education student learning outcomes.
2. Faculty, General Education Steering Committee, and other relevant parties should consider whether or not the CCTDI or the CCTST instruments are aligned well with current (or revised) related student learning outcomes. If neither are aligned, then an alternative assessment(s) that is (are) aligned should be identified.
3. Consider results from the assessments to develop interventions or review and update curriculum to align with areas that need improvement.
4. Determine a timeline to re-collect assessment data related to these student learning outcomes.
5. Increase student participation in future GULL Weeks, particularly in traditionally disproportionately low groups, to increase the likelihood of participant samples that are representative of the entire SU student population, via competitions and marketing to both students as well as faculty that might offer course-embedded incentives for their students that participate.

Contents

Detailed Critical Thinking, Interdependence Among Disciplines, & Intellectual Curiosity Report

Critical Thinking, Interdependence Among Disciplines, & Intellectual Curiosity and Alignment with SU Student Learning Goals

CCTDI Instrument

The California Critical Thinking Disposition Inventory (CCTDI) is an instrument with 75 Likert-style items developed by Insight Assessment to measure the critical thinking mindset. See a CCTDI example item and its alignment with a CCTDI subscale and SU student learning goals in Appendix 1. Details about the instrument can be found at the Insight Assessment CCTDI website (Insight Assessment 2016a) and in the user manual and resource guide (Insight Assessment 2017a). See Table 1 for information on the CCTDI scales.

Table 1. CCTDI scales, based on Insight Assessment CCTDI website (Insight Assessment 2016a) and the user manual (Insight Assessment 2017a)

Scale	Scale Description
Overall Score	The overall measure of critical thinking mindset.
Truth-seeking	"…is the habit of always desiring the best possible understanding of any given situation; it is following reasons and evidence where ever they may lead, even if they lead one to question cherished beliefs. Truth-seekers ask hard, sometimes even frightening questions; they do not ignore relevant details; they strive not to let bias or preconception color their search for knowledge and truth. The opposite of truth-seeking is bias which ignores good reasons and relevant evidence in order not to have to face difficult ideas."
Open-mindedness	"…is the tendency to allow others to voice views with which one may not agree. Open-minded people act with tolerance toward the opinions of others, knowing that often we all hold beliefs which make sense only from our own perspectives. Open-mindedness, as used here, is important for harmony in a pluralistic and complex society where people approach issues from different religious, political, social, family, cultural, and personal backgrounds. The opposite of open-mindedness is intolerance."
Analyticity	"…is the tendency to be alert to what happens next. This is the habit of striving to anticipate both the good and the bad potential consequences or outcomes of situations, choices, proposals, and plans. The opposite of analyticity is being heedless of consequences, not attending to what happens next when one makes choices or accepts ideas uncritically."
Systematicity	"…is the tendency or habit of striving to approach problems in a disciplined, orderly, and systematic way. The habit of being disorganized is the opposite tendency. The person who is strong in systematicity may not know of a given approach, or may not be skilled at using a given strategy of problem solving, but that person has the desire and tendency to try to approach questions and issues in an organized and orderly way."
Confidence in Reasoning	"…is the habitual tendency to trust reflective thinking to solve problems and to make decisions. As with the other attributes measured here, confidence in reasoning applies to individuals and to groups. A family, team, office, community, or society can be trustful of reasoned judgment as the means of solving problems and reaching goals. The opposite habit is mistrust of reasoning, often manifested as aversion to the use of careful reason and reflection when making decisions or deciding what to believe or do."
Inquisitiveness	"…is intellectual curiosity. It is the tendency to want to know things, even if they are not immediately or obviously useful. It is being curious and eager to acquire new knowledge and to learn the explanations for things even when the applications of that new learning are not immediately apparent. The opposite of inquisitiveness is indifference."
Maturity of Judgment	"…the habit of seeing the complexity of issues and yet striving to make timely decisions. A person with maturity of judgment understands that multiple solutions may be acceptable while yet appreciating the need to reach closure at times even in the absence of complete knowledge. The opposite, cognitive immaturity, is imprudent, black-and-white thinking, failing to make timely decisions, stubbornly refusing to change when reasons and evidence would indicate one is mistaken, or revising opinions willy-nilly without good reason for doing so."

CCTST Instrument

The California Critical Thinking Skills Test (CCTST) is an instrument with 34 multiple choice items developed by Insight Assessment. See a CCTST example item and its alignment with a CCTST subscale and SU student learning goals in Appendix 1. Details about the instrument can be found at the Insight Assessment CCTST website (Insight Assessment 2016b) and in the user manual and resource guide (Insight Assessment 2017b). See Table 2 for information on the CCTST scales.

Table 2. CCTST scales, based on Insight Assessment CCTST website (Insight Assessment 2016b) and the user manual (Insight Assessment 2017b)

Scale	Scale Description
Overall Score	"…describes overall strength in using reasoning to form reflective judgments about what to believe or what to do. To score well overall, the test-taker must excel in the sustained, focused and integrated application of core reasoning skills including analysis, interpretation, inference, evaluation, explanation, induction and deduction. The OVERALL score predicts the capacity for success in educational or workplace settings which demand reasoned decision making and thoughtful problem solving."
Analysis	"Analytical reasoning skills enable people to identify assumptions, reasons and claims, and to examine how they interact in the formation of arguments. We use analysis to gather information from charts, graphs, diagrams, spoken language and documents. People with strong analytical skills attend to patterns and to details. They identify the elements of a situation and determine how those parts interact. Strong interpretation skills can support high quality analysis by providing insights into the significance of what a person is saying or what something means."
Inference	"Inference skills enable us to draw conclusions from reasons and evidence. We use inference when we offer thoughtful suggestions and hypotheses. Inference skills indicate the necessary or the very probable consequences of a given set of facts and conditions. Conclusions, hypotheses, recommendations or decisions that are based on faulty analyses, misinformation, bad data or biased evaluations can turn out to be mistaken, even if they have been reached using excellent inference skills."
Evaluation	"Evaluative reasoning skills enable us to assess the credibility of sources of information and the claims they make. We use these skills to determine the strength or weakness of arguments. Applying evaluation skills we can judge the quality of analyses, interpretations, explanations, inferences, options, opinions, beliefs, ideas, proposals, and decisions. Strong explanation skills can support high-quality evaluation by providing the evidence, reasons, methods, criteria, or assumptions behind the claims made and the conclusions reached."
Deduction	"Decision making in precisely defined contexts where rules, operating conditions, core beliefs, values, policies, principles, procedures and terminology completely determine the outcome depends on strong deductive reasoning skills. Deductive reasoning moves with exacting precision from the assumed truth of a set of beliefs to a conclusion which cannot be false if those beliefs are true. Deductive validity is rigorously logical and clear-cut. Deductive validity leaves no room for uncertainty, unless one alters the meanings of words or the grammar of the language."
Induction	"Decision making in contexts of uncertainty relies on inductive reasoning. We use inductive reasoning skills when we draw inferences about what we think is probably true based on analogies, case studies, prior experience, statistical analyses, simulations, hypotheticals, and patterns recognized in familiar objects, events, experiences and behaviors. As long as there is the possibility, however remote, that a highly probable conclusion might be mistaken even though the evidence at hand is unchanged, the reasoning is inductive. Although it does not yield certainty, inductive reasoning can provide a confident basis for solid belief in our conclusions and a reasonable basis for action."
Interpretation	"Interpretative skills are used to determine the precise meaning and significance of a message or signal, whether it is a gesture, sign, set of data, written or spoken words, diagram, icon, chart or graph. Correct interpretation depends on understanding the message in its context and in terms of who sent it, and for what purpose. Interpretation includes clarifying what something or someone means, grouping or categorizing information, and determining the significance of a message."
Explanation	"Explanatory reasoning skills, when exercised prior to making a final decision about what to believe or what to do, enable us to describe the evidence, reasons, methods, assumptions, standards or rationale for those decisions, opinions, beliefs and conclusions. Strong explanatory skills enable people to discover, to test and to articulate the reasons for beliefs, events, actions and decisions."

Alignment of Instruments with the SU Student Learning Goals

SU faculty and UARA staff agreed that the CCTDI and CCTST instruments are aligned with the General Education Critical Thinking, Interdependence Among Disciplines, and Intellectual Curiosity areas and student learning outcomes (Table 3).

Table 3. The SU general education student learning goal, outcomes, and area mapping related to Critical Thinking skills, Interdependence Among Disciplines knowledge, and an Intellectual Curiosity disposition.

Student Learning Goal	Outcome	Area Mapping
1.1. Critical Thinking	1.1.1. Analyze, synthesize, and/or evaluate ideas, concepts, and/or evidence.	IA, IIA, IIB, IIIA, IVA, IVB, IVC, V
	1.1.2. Describe diverse aspects of a discipline using discipline-specific concepts.	IB, IIA, IIB
	1.1.3. Apply appropriate problem-solving strategies to discipline-specific issues.	IIA, IIB, IIIA, IIIB, IVA, IVB, IVC
	1.1.4. Compare and contrast theories within a discipline.	IIA, IIB, IIIB
2.2. Interdependence Among Disciplines	2.2.1. Describe the relationships between at least two different disciplines using principles and/or research of each one.	IA, IIA, IIB, IIIA, IIIB, IVA, IVB, IVC, V
3.3. Intellectual Curiosity	3.3.1. Connect specific examples (e.g., financial literacy, historical currents, social science, STEM, wellness, etc.) to broaden one's life and beyond.	IIA, IIB, IIIB, IVA, IVB, IVC, V
	3.3.2. Expand on course requirements in new and explorative ways.	IIIA

Related to Critical Thinking, Interdependence Among Disciplines, and Intellectual Curiosity, results from these instruments can: provide a benchmark of student outcomes at SU; inform instructional efficacy and possible interventions; evaluate curricular strengths and weaknesses; and continuously monitor and improve student outcomes if we use either instrument for future GULL Week administrations.

Methodology and Sample

Data were collected from volunteer students at SU that self-selected and signed up to participate in various Gaining Understanding as a Lifelong Learner (GULL) Week testing sessions during a week in September 2015 (CCTDI) or February, 2016 (CCTST). GULL Week sessions were open to the entire SU undergraduate student population. The assessments were administered in a proctored computer lab setting and lasted approximately one hour, of which ~35 minutes was dedicated to the CCTDI, whereas ~55 minutes was dedicated to the CCTST. Both instruments were followed by ~5 minutes for a Student Opinion Scale (SOS) Survey (Appendix 2; Sundre & Thelk 2007). The SOS Survey estimates the GULL Week participant's perceived importance of the assessment(s) and effort expended by the participant in completing the assessment(s) (e.g., CCTDI or CCTST).

Some faculty offered incentives (such as extra credit) to participating students, some mentioned GULL Week and encouraged students to participate, and some did not interact with students about GULL Week. The office of University Analysis, Reporting & Assessment (UARA) publicized GULL Week across campus via many avenues. Particularly, competitions between both Schools and Sororities & Fraternities were set up to improve participation.

Based on the manuals for the CCTDI (Insight Assessment 2017a) and the CCTST (Insight Assessment 2017b), the quality and reliability of student responses was defined based on time on task. If a student spent less than 6 minutes on the CCTDI or less than 15 minutes on the CCTDI, then his responses were discarded from the analysis. The students whose data were discarded were then coded as non-test-takers for the respective instruments.

In fall 2015, n = 1359 undergraduates participated in GULL Week and of those n = 515 students completed the CCTDI instrument with quality responses (17.3% and 6.6% of total SU fall 2015 undergraduate enrollment (n = 7849), respectively). In spring 2016, n = 1179 undergraduates participated in GULL Week and of those n = 831 students completed the CCTST instrument with quality responses (15.6% and 11.0% of total SU spring 2016 undergraduate enrollment (n = 7542), respectively). Demographic analyses of the non-CCTDI test-takers (n = 7334; 93.4%) and non-CCTST test-takers (n = 6711; 89.0%) were compared to their respective test-taker groups. This was done to evaluate the extent to which the sample of

test-takers for each instrument was representative of the entire SU undergraduate population during the respective semester. Further analyses within the test-taker groups were performed to evaluate the validity and reliability of the instrument administration at SU as well as to determine whether or not scores on the instrument varied by student characteristic(s). The students with data for one of the Critical Thinking, Interdependence Among Disciplines, and Intellectual Curiosity-aligned instruments and the SOS Survey were analyzed to evaluate student responses on the latter instrument's subscales.

Results

Demographic Comparison of Test-takers vs. Non-test-takers

In general, the demographics of the students that took the CCTDI and CCTST instruments were similar to the non-test-takers, but due to the nature of the sampling method there were groups that were not well represented (Tables 4 – 9).

For CCTDI: Females students (Table 5) were disproportionately high; SU native first time students (Table 6) were disproportionately high; unclassified non-degree undergraduates students were disproportionately low for class level (Table 7); Seidel students were disproportionately high and Henson and undeclared students were disproportionately low (Table 8); and in several cases of student success metrics (i.e., High School GPA, SAT Math, and SU Cumulative GPA), the test-takers of the CCTDI instrument were significantly more successful than the non-test-takers (Table 9). Although it should be considered that for CCTDI the other two success metrics did not reveal any significant differences between the two groups.

For the demographic comparison of the CCTST test-takers and non-test-takers: Asian students were disproportionately high and Unknown/Not specified students were disproportionately low (Table 4); Females students (Table 5) were disproportionately high; SU native first time students (Table 6) were disproportionately high; freshmen were disproportionately high and seniors as well as unclassified non-degree undergraduates were disproportionately low (Table 7); Perdue students were disproportionately high and Seidel and undeclared students were disproportionately low (Table 8); and in two cases of student success metrics (i.e., High School GPA and SU Cumulative GPA) the test-takers of the CCTST were significantly more successful than the non-test-takers (Table 9).

Therefore, in general the samples of CCTDI and CCTST test-takers were only fairly representative of the entire SU undergraduate population during fall 2015 and spring 2016, respectively; however, since some groups were not represented well in these samples any analyses by groups should be interpreted judiciously. In the future, efforts to publicize GULL Week should be targeted more directly to males, transfer students, and students that represent the less successful students (in terms of GPA) as well as continuing previous publicity efforts to ensure even further representative sampling and increased participation.

Table 4. Student Race/Ethnicity Compared between the CCTDI or CCTST Test-takers, Non-test-takers and All SU Undergraduates

Race/Ethnicity	CCTDI Test-taker	CCTDI Non-test-taker	FA15 Total	CCTST Test-taker	CCTST Non-test-taker	SP16 Total
African American	60 (11.7%)	993 (13.5%)	1053 (13.4%)	129 (15.5%)	907 (13.5%)	1036 (13.7%)
American Indian/ Alaska Native	3 (0.6%)	39 (0.5%)	42 (0.5%)	5 (0.6%)	36 (0.5%)	41 (0.5%)
Asian	19 (3.7%)	216 (2.9%)	235 (3.0%)	33 (4.0%)*	183 (2.7%)*	216 (2.9%)
Caucasian	378 (73.4%)	5128 (69.9%)	5506 (70.1%)	556 (66.9%)	4674 (69.6%)	5230 (69.3%)
Hispanic	18 (3.5%)	305 (4.2%)	323 (4.1%)	37 (4.5%)	266 (4.0%)	303 (4.0%)
Native Hawaiian/ Pacific Islander	1 (0.2%)	10 (0.1%)	11 (0.1%)	1 (0.1%)	11 (0.2%)	12 (0.2%)
Non-resident Alien	6 (1.2%)	132 (1.8%)	138 (1.8%)	27 (3.2%)	153 (2.3%)	180 (2.4%)
Two or more races	14 (2.7%)	268 (3.7%)	282 (3.6%)	26 (3.1%)	256 (3.8%)	282 (3.7%)
Unknown/ Not specified	16 (3.1%)	243 (3.3%)	259 (3.3%)	17 (2.0%)*	225 (3.4%)*	242 (3.2%)
Total	515 (100.0%)	7334 (100.0%)	7849 (100.0%)	831 (100.0%)	6711 (100.0%)	7542 (100.0%)

Note. Cell values are counts with percentages reported parenthetically. Significant difference of participation categories between test-takers' and non-test-takers' proportions for each instrument are indicated by an asterisk (*), $p \leq .05$.

Table 5. Student Gender Compared between the CCTDI or CCTST Test-takers, Non-test-takers and All SU Undergraduates

Gender (code)	CCTDI Test-taker	CCTDI Non-test-taker	FA15 Total	CCTST Test-taker	CCTST Non-test-taker	SP16 Total
Male (1)	127 (24.7%)*	3244 (44.3%)*	3371 (43.0%)	268 (32.3%)*	2982 (44.4%)*	3250 (43.1%)
Female (2)	388 (75.3%)*	4085 (55.7%)*	4473 (57.0%)	563 (67.7%)*	3729 (55.6%)*	4292 (56.9%)
Total	515 (100.0%)	7329 (100.0%)	7844 (100.0%)	831 (100.0%)	6711 (100.0%)	7542 (100.0%)

Note. Cell values are counts with percentages reported parenthetically. Significant difference of participation categories between test-takers' and non-test-takers' proportions for each instrument are indicated by an asterisk (*), $p \leq .05$.

Table 6. Student Admit Type, to SU, Compared between the CCTDI or CCTST Test-takers, Non-test-takers and All SU Undergraduates

SU Admit Type (code)	CCTDI Test-taker	CCTDI Non-test-taker	FA15 Total	CCTST Test-taker	CCTST Non-test-taker	SP16 Total
First time student (F)	366 (71.8%)*	4339 (61.1%)*	4705 (61.8%)	583 (70.6%)*	3935 (60.8%)*	4518 (61.9%)
Transfer (T + U)	144 (28.2%)*	2759 (38.9%)*	2903 (38.2%)	243 (29.4%)*	2540 (39.2%)*	2783 (38.1%)
Total	510 (100.0%)	7098 (100.0%)	7608 (100.0%)	826 (100.0%)	6475 (100.0%)	7301 (100.0%)

Note. Cell values are counts with percentages reported parenthetically. Significant difference of participation categories between test-takers' and non-test-takers' proportions for each instrument are indicated by an asterisk (*), $p \leq .05$.

Table 7. Student Undergraduate Class Level Compared between the CCTDI or CCTST Test-takers, Non-test-takers and All SU Undergraduates

Class Level (code)	**CCTDI Test-taker**	**CCTDI Non-test-taker**	**FA15 Total**	**CCTST Test-taker**	**CCTST Non-test-taker**	**SP16 Total**
Freshmen (1)	105 (20.4%)	1490 (20.3%)	1595 (20.3%)	172 (20.7%)*	1062 (15.8%)*	1234 (16.4%)
Sophomores (2)	105 (20.4%)	1579 (21.5%)	1684 (21.5%)	174 (20.9%)	1359 (20.3%)	1533 (20.3%)
Juniors (3)	157 (30.5%)	1970 (26.9%)	2127 (27.1%)	238 (28.6%)	1755 (26.2%)	1993 (26.4%)
Seniors (and +) (4)	143 (28.8%)	1939 (26.4%)	2082 (26.5%)	237 (28.5%)*	2221 (33.1%)*	2458 (32.6%)
Unclassified non-degree undergrads (7)	3 (2.0%)*	356 (4.9%)*	361 (4.6%)	10 (1.2%)*	314 (4.7%)*	324 (4.3%)
Total	153 (100.0%)	7334 (100.0%)	7849 (100.0%)	831 (100.0%)	6711 (100.0%)	7542 (100.0%)

Note. Cell values are counts with percentages reported parenthetically. Significant difference of participation categories between test-takers' and non-test-takers' proportions for each instrument are indicated by an asterisk (*), $p \leq .05$.

Table 8. Student School Enrollment Compared between the CCTDI or CCTST Test-takers, Non-test-takers and All SU Undergraduates

School	**CCTDI Test-taker**	**CCTDI Non-test-taker**	**FA15 Total**	**CCTST Test-taker**	**CCTST Non-test-taker**	**SP16 Total**
Fulton	146 (28.3%)	1856 (25.3%)	2002 (25.5%)	217 (26.1%)	1822 (27.1%)	2039 (27.0%)
Henson	114 (22.1%)*	1928 (26.3%)*	2042 (26.0%)	199 (23.9%)	1578 (23.5%)	1777 (23.6%)
Perdue	106 (20.6%)	1518 (20.7%)	1624 (20.7%)	266 (32.0%)*	1297 (19.3%)*	1563 (20.7%)
Seidel	133 (25.8%)*	1593 (21.3%)*	1685 (21.5%)	129 (15.5%)*	1615 (24.1%)*	1744 (23.1%)
Undeclared	16 (3.1%)*	477 (6.4%)*	496 (6.3%)	20 (2.4%)*	399 (5.9%)*	419 (5.6%)
Total	515 (100.0%)	7462 (100.0%)	7849 (100.0%)	831 (100.0%)	6711 (100.0%)	7542 (100.0%)

Note. Cell values are counts with percentages reported parenthetically. Significant difference of participation categories between test-takers' and non-test-takers' proportions for each instrument are indicated by an asterisk (*), $p \leq .05$.

Table 9. Student Success Metrics Compared between the CCTDI or CCTST Test-takers and Non-test-takers

Success Metric	**CCTDI Test-taker**		**CCTDI Non-test-taker**		**CCTST Test-taker**		**CCTST Non-test-taker**	
	n	**Avg (SD)**	**n**	**Avg (SD)**	**n**	**Avg (SD)**	**n**	**Avg (SD)**
High School GPA	221	3.68 (.47)**	2912	3.55 (.48)**	324	3.63 (.44)*	2579	3.54 (.49)*
SAT Verbal	358	530 (74)	4508	529 (76)	570	525 (78)	4110	529 (75)
SAT Math	358	545 (77)*	4509	536 (78)*	570	538 (81)	4111	535 (77)
SAT Cumulative	358	1075 (133)	4508	1065 (135)	570	1063 (140)	4110	1065 (134)
SU Cumulative GPA	405	3.15 (.57)**	5332	2.97 (.62)**	790	3.12 (.60)*	6178	3.00 (.64)**

Note. Cell values are sample sizes (n) or averages with standard deviation reported parenthetically. Significant difference of participation categories between test-takers' and non-test-takers' average values for each instrument are indicated by an asterisk (*), $p \leq .05$, or two (**), $p \leq .001$.

Validity and Reliability of the Critical Thinking, Interdependence Among Disciplines, & Intellectual Curiosity Instruments' Administrations at SU

CCTDI & CCTST Instrument Common Validity and Reliability Support

The development processes of the CCTDI and CCTST instruments supports their validity and reliability. Also, the results of our administration of the CCTDI and CCTST instruments demonstrated validity. Much of the validity and all of the reliability of the CCTDI and CCTST instruments were described in the manuals (Insight Assessment 2017a, Insight Assessment 2017b, respectively). Content validity was achieved by aligning the items in the instrument with the consensus definition of critical thinking which emerged from the American Philosophical Association (APA) Delphi Study. Communication and cultural validity was established from the input of researchers, professional translators, and native speakers as they developed forms of the instrument across languages. "Test-Retest reliability for all instruments distributed by Insight Assessment meets or exceeds .80 in samples with adequate variance, retested at two weeks post pretest" (Insight Assessment 2017a; Insight Assessment 2017b). Unfortunately, reliability analyses of the SU data incurred additional cost and therefore were not performed. Therefore, the reliability of both tests' administration at SU was not evaluated by the tests' publishers and we could not evaluate their reliability either because the raw data was unavailable from the instruments' publishers. Additional validity and reliability details specific to each individual instrument are detailed below.

CCTDI Instrument SU Administration Validity and Reliability

As per the manual, Likert-style items grouped in scales with demonstrated validity and reliability (Insight Assessment 2017a), specifically, in terms of reliability,

> *Internal consistency reliability for the seven individual scales in the initial CCTDI pilot sample ranged from .71 to .80, with the alpha for the overall instrument reaching or exceeding .91. Strong values have been observed consistently in samples collected over the past 15 years (ranging from .60 to .78 on the scales and .90 or above for the overall measure).*

Based on the SU student scores in fall 2015, criterion and construct validity were supported because students' overall score on this instrument had positive correlations with the SU students' related measures of SAT Verbal scores, $r = .363$ ($p < .001$), SAT total scores, $r = .304$ ($p < .001$), and SAT Math scores, $r = .169$ ($p < .05$). Correlation coefficients ≥ .1 ≤ .3 are evidence of small effect sizes and ≥ .3 ≤ .5 are evidence of medium effect sizes (Field 2013).

CCTST Instrument SU Administration Validity and Reliability

Scale validity of the instrument was evaluated over the last 40 years to determine alignment of items in the item pool to scale, with acceptable factor loadings ranging from .300 - .770 (Insight Assessment 2017b). Item analysis has also been performed over the last 40 years to determine the item difficulty of items in the item pool and in conjunction with acceptable Kuder-Richardson (KR-20) scores, "The OVERALL Scores of all versions of the reasoning skills tests meet or exceed this .70 criterion in the validation samples, and in large model population samples," support the reliability of the CCTST. Furthermore, criterion and construct validity based on both increases on the test following critical thinking interventions as well as "high correlations with standardized tests of college-level preparedness in higher-order reasoning have been demonstrated (GRE Total Score: Pearson $r = .719$, $p<.001$; GRE Analytic $r = .708$, $p<.001$; GRE Verbal $r = .716$, $p<.001$; GRE Quantitative, $r = .582$, $p<.001$)" (Insight Assessment 2017b). Based on the SU student scores in spring 2015, criterion and construct validity were supported because students' overall score on this instrument had positive correlations with the related measures of SAT Verbal scores, $r = .498$, Math scores, $r = .495$, and total scores, $r = .560$ (all $p < .001$) as well as High School GPA, $r = .178$ and SU Cumulative GPA, $r = .133$ (all $p < .05$). Correlation coefficients ≥ .1 ≤ .3 are evidence of small effect sizes; ≥ .3 ≤ .5 are evidence of medium effect sizes; and ≥ .5 are evidence of large effect sizes (Field 2013).

SU Student Scores on Instruments

On average, the students that participated (n = 515) had an overall score of 295 (SD = 27) with a range of 214 to 369 on the CCTDI instrument. The possible overall CCTDI score range is 70 – 420, which is additive of the seven attribute scales whose possible scores ranges are 10 – 60. For qualitative descriptions of the numerical scores, see Table 10. Therefore, in general, our students were "Positive" for critical thinking mindsets although the range was from "Inconsistent/Ambivalent" to

"Strong Positive." For most of the scales, the SU students were "Positive" for the aligned attribute, but for both Truth-seeking and Systematicity there is evidence for a need for improvement (Table 11).

Table 10. Recommended CCTDI Mindset Attribute Score Interpretations (Insight Assessment 2017a)

Overall Numerical Score Range	Attribute Scales Numerical Score Range	Qualitative Category	Recommended Attribute Interpretation
350 – 420	50 – 60	Strong Positive	Scores in this range indicate that the attribute or attitude is a positive habit of mind and likely to factor into the individual's approach to all higher order thinking (reflective problem definition and problem solving), particularly when the situation is of high consequence.
280 – 349	40 – 49	Positive	These scores indicate consistent endorsement and valuation of the attitude or attribute being measured. Individuals scoring in this range typically demonstrate this mindset attribute with reliability.
210 – 279	30 – 39	Inconsistent/ Ambivalent	These scores are indicative of ambivalent or inconsistent endorsement of the attitude or attribute being measured. Individual test takers are frequently seen to move from this score category to a higher range score as a result of completing an educational or training program aimed at developing this mindset attribute. More rarely, such efforts also result in some individuals regressing to a more adverse or hostile range at second assessment. This movement toward aversion or hostility is consistent with other observations of attitude formation. Having the habit of mind of engaging life and work problems with one's higher order thinking skills in order to determine what to believe and what to do (critical thinking), requires courage, persistence, honesty, and organization. It is reasonable that some individuals may reject the challenge or find this path too fearful or difficult. Many of these fall into non-reflective strategies to address life and work problems.
140 – 209	20 – 29	Negative	These scores are indicative of poor valuation or aversion toward the attribute being measured. Scores in this range indicate negative habits of mind, and approaches to reasoning and problem solving that are likely to limit the quality of thinking and decision-making, and diminish the fair-minded and willing engagement of problems in many situations. While movement from this level of score to one demonstrating strength on the attribute has been seen in some individuals after participation in an educational opportunity aimed at the specific attribute in question, it is more common to see an individual express at most an ambivalent valuation of the measure at the second assessment date.
70 – 139	10 – 19	Strong Negative	These scores are indicative of strong negativity or hostility toward the attribute being measured. Scores in this range indicate that the individual may systematically avoid situations requiring thinking and decision making or work to disrupt or destroy decision processes where others seek to make strong judgments. Movement from this score range to positive score ranges after participation in an educational opportunity aimed at strengthening the specific attribute in question is rarely observed.

Table 11. Summary of SU Students' scores on the attribute scales of the CCTDI

Scale	Avg (SD) *Qualitative Category*	Percent of Students (n = 515)		
		Proficient	Developing	Needs Improvement
Overall	295 (27) *Positive*	70.3%	29.7%	-
Truth-seeking	35.2 (5.7) *Inconsistent/ Ambivalent*	19.6%	64.9%	15.5%
Open-mindedness	43.1 (5.5) *Positive*	72.8%	26.4%	.8%
Analyticity	44.6 (4.8) *Positive*	87.8%	11.8%	.4%
Systematicity	39.7 (6.2) *Inconsistent/ Ambivalent*	48.9%	45.4%	5.6%
Confidence in Reasoning	44.8 (6.0) *Positive*	82.7%	16.5%	.8%
Inquisitiveness	46.2 (5.9) *Positive*	87.4%	12.0%	.6%
Maturity of Judgment	41.4 (6.3) *Positive*	66.0%	30.1%	3.9%

Note. Proficient = students with scores ranging from Positive to Strong Positive; Developing = students with Inconsistent/Ambivalent scores; Needs Improvement = students with scores ranging from Strong Negative to Negative (see Table 10 for numeric value ranges). Highlighted values denote high percentages of students in the "Needs Improvement" category.

On average for the CCTST instrument, the students that participated (n = 831) had an overall score of 72.2 (SD = 6.0) with a range of 55 to 94. The possible overall CCTST score, as well as the scales' scores, range is 50 – 100, which is normalized. For qualitative descriptions of the numerical scores, see Table 12. Therefore, in general, our students were "Moderate" for critical thinking skills although the range was from "Not Manifested" to "Superior." For all of the scales, the SU students were "Moderate" for each aligned skill, which suggests improvement is needed across all scales. However, there is evidence for a targeted need for improvement in the Evaluation, Deduction, and Explanation skills since there were higher percentages of students in the "Needs Improvement" category (Table 13). Based on the distribution of the overall score percentiles for the test takers in our SU group, as compared to an aggregate sample of CCTST Four Year College Students (national undergraduate students from four year colleges and universities), the average percentile of our group of students in this sample is 35.4 (SD = 23.6).

Table 12. Recommended CCTST 100 point Performance Assessments Score Interpretations (Insight Assessment 2017b)

Numerical Score Range	Qualitative Category	Recommended Performance Assessments Interpretation
86 – 100	Superior	This result indicates critical thinking skill that is superior to the vast majority of test takers. Skills at the superior level are consistent with the potential for more advanced learning and leadership.
79 – 85	Strong	This result is consistent with the potential for academic success and career development.
70 – 78	Moderate	This result indicates the potential for skills-related challenges when engaged in reflective problem-solving and reflective decision-making associated with learning or employee development.
63 – 69	Weak	This result is predictive of difficulties with educational and employment related demands for reflective problem solving and reflective decision making.
50 – 62	Not Manifested	This result is consistent with possible insufficient test taker effort, cognitive fatigue, or possible reading or language comprehension issues.

Table 13. Summary of SU Students' scores on the attribute scales of the CCTST

Scale	Avg (SD) *Qualitative Category*	Percent of Students (n = 831)		
		Proficient	Developing	Needs Improvement
Overall	72.2 (6.0) *Moderate*	15.2%	48.4%	36.5%
Analysis	73.3 (7.6) *Moderate*	27.7%	50.3%	22.0%
Inference	74.9 (6.2) *Moderate*	23.1%	52.3%	24.5%
Evaluation	70.5 (7.8) *Moderate*	18.2%	36.8%	45.0%
Deduction	71.8 (6.3) *Moderate*	14.8%	43.2%	42.0%
Induction	76.6 (6.8) *Moderate*	42.1%	41.0%	16.8%
Interpretation	78.8 (9.1) *Moderate*	55.8%	23.9%	20.2%
Explanation	72.0 (9.6) *Moderate*	27.6%	25.3%	47.2%

Note. Proficient = students with scores ranging from Strong to Superior; Developing = students with Moderate scores; Needs Improvement = students with scores ranging from Not Manifested to Weak (see Table 12 for numeric value ranges). Highlighted values denote high percentages of students in the "Needs Improvement" category.

On average, there was statistical difference in the overall score on the CCTDI instrument for SU native first time students as compared to transfer students (Table 14), but there was no significant difference between those two groups for the overall score on the CCTST instrument. SU native first time students had lower CCTDI overall scores as compared to transfer students and the difference, -7, was significant $t(508) = -2.58$, $p < .05$. After centering age, using regression analysis, we analyzed the effect of age as a moderator on the relationship between admit type and the CCTDI overall score. There is no significant interaction, $b = 1.743$, 95% CI [-.573, 4.058], $t = 1.478$, $p > .05$, to indicate that age is a moderator of the relationship between admit type and the CCTDI overall score.

SU native first time students had similar CCTST overall scores as compared to transfer students and the difference, .2, was not significant $t(824) = .42$, $p > .05$.

Table 14. Student Admit Type, to SU, Average Overall Scores on the CCTDI and CCTST Instruments

SU Admit Type (code)	CCTDI			CCTST		
	n	Score	SD	n	Score	SD
First time student (F)	366	293	27	583	72.3	5.7
Transfer (T + U)	144	300	27	243	72.1	6.6

As SU's students' class level increased from freshman to senior, so too did the average overall score on the CCTDI instrument (Table 15). However, there was no significant difference between class level groups for the overall score on the CCTDI ($F(4, 510) = 1.394$, $p > .05$). When student class level was analyzed as a fixed factor and student SU admit type was included as a covariate, although the covariate ($F(1, 505) = 3.930$, $p < .05$) has a significant effect, the fixed factor after controlling for the effect of the covariate ($F(4, 505) = .906$, $p > .05$) was not significantly related to the participant's overall score on the CCTDI.

Similar results were found for the CCTST average overall score (Table 15), except there was a significant difference between class level groups with a small effect size ($F(4,65.661) = 8.504$, $p < .001$, $r = .208$). Post hoc comparisons, via the Tukey HSD test, were used to identify which class levels' average scores were significantly different. Tests revealed significant pairwise differences between the average scores of Freshmen compared to Juniors, Seniors, and Unclassified undergraduates as well as between Sophomores and Seniors and Unclassified Undergraduates. Other relationships between class level groups were

not significantly different. Also, when student class level was analyzed as a fixed factor and student SU admit type was included as a covariate, both the covariate ($F(1, 820) = 7.134$, $p < .01$) and the fixed factor after controlling for the effect of the covariate ($F(4, 820) = 10.809$, $p < .001$) were significantly related to the participant's CCTST overall score. After centering age, using regression analysis, we analyzed the effect of age as a moderator on the relationship between class level and the CCSTS overall score. Although there is no interaction, $b = .126$, 95% CI [-.092, .344], $t = 1.134$, $p > .05$ to indicate that age is a moderator of the relationship between class level and the CCTST overall score.

Table 15. Student Undergraduate Class Level Average Overall Scores on the CCTDI and CCTST Instruments

Class Level (code)	**CCTDI**			**CCTST**		
	n	**Score**	**SD**	**n**	**Score**	**SD**
Freshmen (1)	105	291	29	172	70.5	5.2
Sophomores (2)	105	292	26	174	71.4	6.3
Juniors (3)	157	296	25	238	72.7	5.4
Seniors (and +) (4)	143	298	29	237	73.4	6.2
Unclassified non-degree undergrads (7)	5	299	28	10	77.3	10.4

Student performance by SU School is listed in Table 16. There was a significant difference in the CCTDI overall score based on enrollment in School at SU, but the difference in average scores between groups was quite small based on effect size value interpretation ($F(4, 510) = 3.314$, $p < .05$, $r = .16$). Post hoc comparisons, via the Tukey HSD test, were used to identify which Schools' average scores were significantly different. Tests revealed significant pairwise differences between the average scores of Fulton School's students as compared to Seidel School's students, $p < .05$, where the former had higher scores. Students from Henson or Perdue Schools or that are undeclared do not significantly differ from the other groups, $p > .05$. In contrast, there was no significant difference in the CCTST overall score based on enrollment in School at SU, ($F(4,826) = 2.032$, $p > .05$).

Table 16. Student School Enrollment Average Overall Scores on the CCTDI and CCTST Instruments

School	**CCTDI**			**CCTST**		
	n	**Score**	**SD**	**n**	**Score**	**SD**
Fulton	146	299*	26	217	72.0	6.1
Henson	114	296	28	199	73.1	6.5
Perdue	106	296	27	266	72.3	5.7
Seidel	133	288*	27	129	71.3	5.5
Undeclared	16	290	27	20	72.2	6.6

Note. Significant difference of categories' average scores for each instrument are indicated by an asterisk (*), $p \leq .05$.

Although not presented here, student performance by primary major is available upon request to programs or Departments when at least 30 students in that major participated in this instrument's administration. These data can be used for informal review and improvement efforts, or for more formal program review and improvement efforts such as Academic Program Review required reporting related to assessment of program student learning outcomes aligned with this instrument, when applicable.

Critical Thinking, Interdependence Among Disciplines & Intellectual Curiosity Instruments and SOS Survey Student Responses

The CCTDI and CCTST test-takers also took the SOS Survey (CCTDI n = 507, CCTST n = 818; Table 17). We were able to evaluate the reliability of both subscales for each test within the SOS Survey. The *Importance* subscale, which addresses the extent to which the student thought it was important to do well on the test each test-taker participated in, demonstrated reliability (CCTDI $\alpha = .789$, CCTST $\alpha = .741$). Similarly, the *Effort* subscale, which addresses the extent to which the student fully engaged in effortful behavior on the test s/he participated in, demonstrated reliability (CCTDI $\alpha = .840$, CCTST $\alpha = .779$). The validity of the instrument is discussed in the SOS Survey Test Manual (Sundre & Thelk 2007). The 10 items, five in

each subscale, are measured in a 1 to 5 scale, where 1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, and 5 = Strongly Agree. There are four items that are negatively worded, and their scores were reverse coded prior to analysis.

In general, students that took the CCTDI were fairly "Neutral" in their responses for both the *Importance* and *Effort* subscales. For *Importance,* this indicates that students thought that their scores on the instrument would not affect them strongly in either a negative or positive way. For *Effort*, it indicates that students put in a moderate effort towards completing the instrument. The two subscale scores had a positive correlation with one another, r = .481 (p <.01; medium effect size) and both subscales were also negatively correlated with the CCTDI instrument overall score (*Importance*, r = -.317 (p < .01; medium effect size); *Effort*, r = -.356 (p < .01; medium effect size)). The latter two correlations of SOS subscales with the CCTDI overall score indicate that the students that self-reported exerting more effort or that thought that this instrument was important also had lower overall scores than those who did not self-report exerting as much effort or that thought this instrument was not important, with medium effect sizes.

In contrast, for the CCTST instrument, in general students reported that they "Agree[d]" with the statements. For *Importance,* this indicates that students thought that their scores on the CCTST instrument would affect them in either a negative or positive way. For *Effort*, it indicates that students put in a concerted effort towards completing the instrument. The CCTST overall total score was significantly correlated with the *Importance* subscale score, r = .116, and the *Effort* subscale score, r = .248; the *Importance* and *Effort* subscale scores were also significantly correlated with one another, r = .421 (all p values ≤ .001). This indicates that the students that self-reported exerting more effort or that thought that this instrument was important also scored higher on the instrument score than those who did not self-report exerting as much effort or that thought this instrument was not important, with low effect sizes. Students that self-reported this instrument was more important also exerted more effort, with a medium effect size.

Table 17. Student Opinion Scale (SOS) Survey subscales' administrative results for the students that also participated in the CCTDI or CCTST instruments' administration.

Instrument	SOS Subscale	Number of Items	Reliability (α)	n	Average Score (out of 25)	SD
CCTDI	**Importance**	5	.789	506	13.5	3.8
	Effort	5	.840	507	10.4	3.4
CCTST	**Importance**	5	.741	818	16.3	3.6
	Effort	5	.779	818	18.3	3.3

Discussion

Based on the results presented here it seems that there is room for improvement in student learning outcomes related to Critical Thinking, Interdependence Among Disciplines, and Intellectual Curiosity at SU. Several action items are suggested below towards this end.

1. Faculty, the General Education Steering Committee, and other relevant stakeholders should evaluate the need to revise the current Critical Thinking, Interdependence Among Disciplines, and Intellectual Curiosity General Education Areas' student learning outcomes. Do the current student learning outcomes align with our expectations of students' knowledge and skills related to those areas that should be achieved while they are at SU? Is the language clear? Are they assessable? These should be targeted at the institutional level, but other levels of student learning outcomes related to Critical Thinking, Interdependence Among Disciplines, and Intellectual Curiosity may be generated as well to address program, or course-level assessment needs.
2. Based on discussions and decisions related to Action Items #1, relevant parties such as faculty and the General Education Steering Committee should consider whether or not the CCTDI or CCTST instruments are aligned well with the current (or revised) SU Critical Thinking, Interdependence Among Disciplines, and Intellectual Curiosity General Education Areas' student learning outcomes. If neither are aligned, then an alternative assessment(s) that is (are) aligned should be identified (e.g., ETS HEIghten-Critical Thinking instrument).
3. Relevant stakeholders at SU should consider the results from the CCTDI and CCTST instruments to develop interventions or review and update curricula to align with areas that need improvement. In particular, focus should be on the scales that, based on this sample of SU students, require the most improvement in Critical Thinking mindsets (Truth-seeking and Systematicity) and skills (Evaluation, Deduction, and Explanation).
4. Based on discussions and decisions related to Action Items #1-3, a timeline for re-assessment of the SU Critical Thinking, Interdependence Among Disciplines, and Intellectual Curiosity General Education Areas' student learning outcomes should be proposed. This will allow an analysis of whether or not there is change in student learning outcomes based upon either a change in assessment or instructional or curricular interventions.
5. Attempt to increase student participation in future GULL Weeks, particularly in traditionally disproportionately low groups, to increase the likelihood of participant samples that are representative of the entire SU student population. This can be done via efforts that have occurred in the past, such as competitions and marketing to both students as well as faculty that might offer course-embedded incentives for their students that participate. However, new ways to incentivize participation of traditionally disproportionately low groups should also be identified and implemented.

References

Field, A.P. 2013. Discovering statistics using IBM SPSS Statistics (4th ed.). Thousand Oaks, CA: Sage Publications, Inc.

Insight Assessment. 2017a. "CCTDI User Manual and Resource Guide." The California Academic Press. San Jose, CA.

Insight Assessment. 2017b. "CCTST User Manual and Resource Guide." The California Academic Press. San Jose, CA.

Insight Assessment. 2016a. California Critical Thinking Disposition Inventory (CCTDI). Retrieved November 29, 2016, from http://www.insightassessment.com/Products/Products-Summary/Critical-Thinking-Attributes-Tests/California-Critical-Thinking-Disposition-Inventory-CCTDI

Insight Assessment. 2016b. California Critical Thinking Skills Test (CCTST). Retrieved November 29, 2016, from http://www.insightassessment.com/Products/Products-Summary/Critical-Thinking-Skills-Tests/California-Critical-Thinking-Skills-Test-CCTST

Insight Assessment. 2016c. Sample Items for Measuring Thinking Attributes. Retrieved November 30, 2016, from http://www.insightassessment.com/Resources/Sample-Items-for-Measuring-Thinking-Attributes/(language)/eng-US

Insight Assessment. 2016d. Sample Thinking Skills Questions. Retrieved November 30, 2016, from http://www.insightassessment.com/Resources/node_1487/(language)/eng-US

Sundre, D.L., & Thelk, A.D. 2007. "The Student Opinion Scale (SOS): A Measure of Examinee Motivation: Test Manual." James Madison University, The Center for Assessment & Research Studies.

Appendices

Appendix 1. Example items from the CCTDI and CCTST Instruments

CCTDI Example Items.
NOTE: These are not items on the instrument, but are representative of items on the instrument.

Consider the following statements about beliefs, opinions, values, and preferences. Decide whether you agree or disagree with each one. Remember that since you are being asked about your own beliefs, opinions, values, and preferences, there really is no "right" or "wrong" response. The answer is whatever you say it is for you.

Use the following choices to express your view:
6 = Agree Strongly
5 = Agree
4 = Agree Marginally
3 = Disagree Marginally
2 = Disagree
1 = Disagree Strongly

1. People say I ask challenging questions.
2. I won't let what scientists say weaken my core beliefs.
3. I prefer jobs where the supervisor says exactly what to do, and exactly when and how to do it.
4. It's important to me to figure out what people really mean by what they say.
5. Don't kid yourself, changing your mind is a sign of weakness.
6. I always do better in jobs where I'm expected to think things out for myself.
7. I hate it when people just shout their opinions without letting others give their views too.
8. There is never any good reason for believing one thing rather than another.
9. Being organized about your plans and projects is way over-rated.
10. Don't try to think ahead because it is impossible to know exactly what the future holds.

Correct Response: n/a

Alignment:
CCTDI Scales: There is no indication of item alignment with instrument scales, but see Table 1 for scale descriptions.
SU General Education Areas of Critical Thinking, Interdependence Among Disciplines, & Intellectual Curiosity: see Table 3 for descriptions.

CCTST Example Items.
NOTE: These are not items on the instrument, but are representative of items on the instrument.

For example items 1, 2, and 3 please consider this information:
A scientific study compared two matched groups of college women. The women in both groups were presented with information about the benefits of a healthy diet and regular exercise. The women in one group were paired up with one another and encouraged to work as two-person teams to help each other stick with the recommended healthy regimen of smart eating and regular vigorous exercise. The women in the other group were encouraged to use the same recommended regimen, but they were also advised to work at it individually, rather than with a partner or teammate. After 50 days the

physical health and the well-being of all the women in both groups were evaluated. On average the women in the first group (with teammates) showed a 26 point improvement in measures of cardiopulmonary capacity, body strength, body fat reduction, and sense of well-being. On average the women in the other group (encouraged to work as individuals) showed a 17 point improvement on those same measures. Using statistical analyses the researchers determined that the probability that a difference of this size had occurred by chance was less than one in 1000.

Example Item 1: If true, these research findings would tend to support which of the following assertions?

A. A college woman cannot achieve optimal health functioning without a teammate.
B. Universities should require all students living in campus residence halls to participate in a health regime of smart eating and regular vigorous exercise.
C. A healthy diet will cause one to have better mental health and physical strength.
D. This research study was funded by a corporation that makes exercise apparel.
E. A regimen of smart eating and regular exercise is related to better health.

Example Item 2: If the information given in the case above were true, which of the following hypotheses would not need to be ruled out in order to confidently claim that for the majority of young adults a regimen of smart eating and regular vigorous exercise will result in significant improvements in one's overall health?

A. This study was about women, the findings cannot be generalized to include men.
B. Since the study began to solicit willing participants before the Research Ethics Review Committee of the college gave the research project its formal approval to gather data, the findings are invalid.
C. Some women in the study over-reported their compliance with the eating and exercise regimen, which led the researchers to underestimate the full impact of the regimen.
D. Since many of those studied described themselves as overweight or out of shape when the study began, a similar regimen will not benefit people who are healthier to start with.
E. The performance tests used to evaluate the health and well-being of females may not be appropriate for evaluating the health and well-being of males.

Example Item 3: Consider the claim, "Working with a teammate or partners on a health regimen is better than working individually." Which of the following additional pieces of information would not weaken that claim?

A. Most of the women in the group that was encouraged to work individually actually worked with friends and partners who were not part of the study.
B. Most of the pairings and teams created in the first group (with teammates) fell apart after a few days and the women in that group actually worked individually.
C. There was something about the women in the first group (with teammates) that the researchers overlooked, thus invalidating the intended matching of the two groups.
D. Men are more likely to work alone, so any recommendation that men find a teammate or partner to support them in sticking with the regimen will be ignored.
E. The study was undertaken when there were no exams or major projects due, thus the results about working with a teammate do not apply to more stressful times of the year.

Correct Responses: not provided

Alignment:
CCTST Scales: There is no indication of item alignment with instrument scales, but see Table 2 for scale descriptions.
SU General Education Areas of Critical Thinking, Interdependence Among Disciplines, & Intellectual Curiosity: see Table 3 for descriptions.

Appendix 2. Student Opinion Scale (SOS) Survey (Sundre & Thelk 2007)

Item	Item Text	Subscale
1	Doing well on these tests was important to me.	Importance
2	I engaged in good effort throughout these tests.	Effort
3*	I am not curious about how I did on these tests.	Importance
4*	I am not concerned about the scores I receive on these tests.	Importance
5	These were important tests to me.	Importance
6	I gave my best effort on these tests.	Effort
7*	While taking these tests, I could have worked harder on them.	Effort
8	I would like to know how well I did on these tests.	Importance
9*	I did not give these tests my full attention while completing them.	Effort
10	While taking these tests, I was able to persist to completion of the tasks.	Effort

* Denotes items that are reversed prior to scoring.