STUDENT ASSESSMENT SUMMARY
Jule Dee Scarborough, Ph.D.
(See Portfolio Sections B.9.a,b,c,d, and B.9.e.1-5; also, A.5 and A.7)

The goals of this program component were to diversify assessment so students had a wider range of types of assessments and more opportunities to provide evidence of learning – a multifaceted student assessment plan for the course. Another goal was to improve the quality of the midterm and final exams by designing them to incorporate a wider range of types of items and improving the problems to be solved, adding more specific grading criteria. Also it was important to measure the capability of the assessments to engage students at the upper levels of Bloom's Cognitive Dimension, as tests and other assessments can be designed at a much greater quality level when using Bloom's Cognitive Dimension to determine if the questions stimulate higher level thinking to reveal a deeper understanding of the concepts, principles, knowledge or skills. Finally, there was an additional goal of adding three formal performance assessments, with the corresponding rubrics for scoring performance of each, to the assessment plan. The above tasks were all accomplished, and once completed, the professors analyzed each assessment procedure against Bloom's Cognitive Dimension.

Test Analysis and Development
During the course analysis program component, the professors performed analyses of their midterm and final exams to use for diagnostic purposes – for the improvement of the tests and instruction. The test analyses were used to design and develop new midterms and final examinations for the 2006 course to be taught during the experimental research semester. Therefore, regarding the timeline, the test analyses were performed after the course, teaching, and learning analyses were completed. (See the GAPS Analysis in Portfolio Section B.5 for those results.)

After identifying the strengths and weaknesses of the tests used in the 2005 course, the professors then developed new Student Learning Outcomes for the 2006 course, as described in the previous Outcomes section. In the "Reversed and Intentional Instructional Design" process, the Outcomes define the "what" is to be learned and then what type of assessments are to be used for measuring student learning is determined. This answers the question "what is acceptable evidence of learning?" The professors designed and developed new objective midterm and final examinations; some professors included subjective items as well. (Refer to the section following this summary to read the analysis results for the 2005 tests as compared to the new 2006 tests, Portfolio Section B.8.2.)

Briefly, many of the professors included mathematical problems on their 2005 tests, some had objective items, and a few had no objective items. The range of types of tests varied across professors. The strength of some tests might be the problems, except that there were no grading criteria or rubrics with standards or criteria of performance. Therefore, the grading of problems was not perceived as consistent nor could it be determined exactly what professors were seeking beyond the "right" answer. Some professors did give points along the way for "right" partial elements of solutions; others did not. The grading was far too unspecified and far too subjective to really mean much. Also the point distribution or grade distribution was not formulated based upon objective criteria. Finally, for most of the courses, the tests seemed to be the primary student assessments, with very few (possibly homework) other opportunities to further assess student learning. Several did include design projects, but these also were missing formal grading formula, criteria, or rubrics. Once again the grading was perceived as far too subjective, and the potential inconsistencies and lack of formal and predetermined objectivity left grading ambiguous. Professors lacked a multifaceted assessment plan for their courses where students were assessed for learning using a variety of types of assessments. They also provided few opportunities for students to demonstrate evidence of learning, meaning that, there were not very many assessments. Students are at a real disadvantage where there are very few assessments.

Once the 2005 test analyses were completed and we determined that there were no performance assessments in any of the courses, the professors designed and developed new midterm and final exams for the 2006 course. Using the results of those analyses diagnostically and the new 2006 student learning outcomes, the professors developed a Table of Specifications to guide their creation of a new midterm and final exam for the 2006 course. They each developed an objective test item bank of multiple questions for each student learning outcome. Once the objective test items were developed, they chose items for each exam, midterm and final, and assembled the tests. If they preferred to include problems to solve, those were added as well. The program leaders provided feedback throughout the entire analysis and development process. To further ensure that the test items and assembled tests actually measured knowledge or skills inherent in an outcome, the professors mapped the outcomes to the corresponding tests and specific items. This helped them realize where they needed more items or a different type of item and, especially, where there were gaps in the measurement of critical outcomes. (See worksheet below and other examples in B.9.2.e.1-5.) Although the tests were not perfect, they were greatly improved. (See Portfolio Section B.9.a for the comparison for differences between the 2005 and 2006 tests; see Program Description, Portfolio Section A.7, for further information on the Test Analysis and Development program components.)

An example of the outcome to test and test item analysis is copied here.

Table B.9.1: IENG 475 - Decision Analysis – Regina Rahn

<table>
<tr><th colspan="4">Student Learning Objectives and Outcomes</th><th>Assessments:
Test Alignments
Midterm & Final</th></tr>
<tr><td></td><td>Student Learning Objectives/Outcomes-Major</td><td colspan="2">Student learning objectives - minor</td><td>Corresponding Tests and Test Items</td></tr>
<tr><td>I</td><td>To learn to use a specific set of analytical tools for technical decision making under uncertainty.</td><td>A</td><td>Students will be able to construct/create a decision tree to aid in determining the best course of action for a given set of circumstances
1. To define the states of nature of the system, process, or situation
2. To develop the branch structure of the tree
a. To identify decision nodes; what are the items the decision maker chooses
b. To identify the chance nodes; the events that occur by chance with a given probability
c. To draw the arcs, which define the sequences and relationships between nodes
3. To identify the outcomes
a. To define the choices for a decision node
b. To define the possible outcomes of a chance node, which are a set of mutually exclusive outcomes
c. To define the “consequence,” or the final outcome of a branch
4. To solve for the expected value of the decision tree (EV, EMV)
a. To construct the joint, conditional, and marginal probabilities
b. To calculate all branch probabilities of the tree
i. To apply Baye’s Theorem
ii. To implement the inverse tree structure technique
5. To find and compare the expected value of both sample and perfect information (EVPI, EVSI)
a. To construct the decision trees to calculate EVPI and EVSI
b. To evaluate the relevance and importance of the values obtained for EVPI and EVSI to the decision process</td><td>HW #1 (A)

HW #2 (A)

HW #3 (B,C)

HW #4 (D)

HW #5 (E)

Performance Task #1 (A)

Midterm # 1-18 (A)

Midterm # 19-21(B)

Midterm # 22-30 (C)

Final # 19-22 (A)

Final # 6-10, 23-25 (C)

Final # 1-5 (D)

Final # 11-18 (E)</td></tr>
</table>

Performance Assessment (See B.9.b and c; also, A.5 and A.7)
The second major program component related to student assessment was Performance Assessment. Each professor developed three performance tasks and corresponding rubrics. One performance task/rubric corresponded with the midterm examination and a second corresponded with the final exam. The third performance task/rubric use and focus was determined by the professors. Why design the two performance tasks to correspond to the midterm and final examinations – because the underlying premise that we chose to accept was that "most" or "typical" tests, at best, provided evidence of what students may "know about" or "know" at the lower levels of Bloom's Cognitive Dimension (memorize or limited comprehension). We do acknowledge that good tests may achieve these Bloom's levels, but tests do not "usually" provide the opportunity for students to apply, analyze, evaluate, synthesize or create. Furthermore, if well designed and developed problems are used, either for students to solve with established performance criteria for judging performance or to have to figure out before responding to a selection of provided responses, then these well designed and developed tests can accomplish the goal of providing high quality evidence of learning. But developing such tests really requires knowledge and skill in test analysis and development, as well as a great deal of time. When considering that most engineers and technology professors, as well as most of the university population of professors, have little to no background in educational or learning theory or in student assessment, then the reality is that most tests may not measure anything of significance very well or provide any worthy evidence of learning. Therefore, if one accepts that well designed and developed performance tasks (authentic and real world), with high quality and well defined performance criteria through the use of rubrics, can provide a greater opportunity for students to provide evidence of what they can "do" with the knowledge and skills, then performance assessment extends the evidence of learning possibilities. However, performance tasks and objective tests, even with good problems, can measure some similar things, but also very different knowledge and skills. Thus, we wanted the performance tasks to relate, but also to extend the evidence learning to include "doing" or "performing." We had hoped to consider the difference and similarities between the two types of measures, a good study in itself if formally executed. But we realized that the professors did not have background enough in assessment or the theory needed to design such research as this point in their experience. However, this was our way of introducing performance tasks/rubrics while also asking the professors to improve their tests. Some of the professors did not have objective test items at all on their tests; there were just subjective problems to solve without any established performance criteria. We wanted them to understand that if they were going to use solving a problem as a high quality performance, then they had to go the next step and develop a rubric that revealed the performance criteria for solving the problem, making it clear to both the student and the professor what the standards and criteria of performance could be. This also helped them to understand that they needed to grade more than just the end result or answer, that problem solving was a process that resulted in an answer.

For this first venture into performance assessment, we required three "complex" performance tasks, where each performance task was actually a cluster of several performances, was authentic, and reflected real world tasks, and the rubrics revealed levels of possible standards for performance and the criteria for each standard. Each professor developed the three performance tasks with corresponding rubrics.

Both professors and students reported liking performance assessment. The first ones were somewhat simple, but that will change as the professors continue to modify them and become more skilled at developing the tasks and rubrics. All professors reported that they will continue to use performance tasks and rubrics. This type of assessment greatly improved their overall course-student learning assessment plans, further diversifying the types of assessment for the course. Performance tasks also added more opportunities to provide evidence of learning. (See the Assessment Plan Maps included below, example and professors'.) Finally, each professor studied his/her overall tests, performance tasks, and any other types of assessment they included (e.g., design projects, collaborative projects, etc.) and analyzed them for achieving the higher levels on Bloom's Cognitive Dimension. There was a greatly improved range of Bloom's upper levels. For most of the professors, the tests achieved well to the application level, with three midterms and finals achieving the analysis and synthesis levels. Clearly, progressing to the inclusion of formal performance tasks moved the assessment plans to the higher levels of Bloom's Cognitive Dimension.

The Chart has been copied into this document below for reader convenience.

Table B.9.2: Bloom's Learning Taxonomy –Assessment Analysis Chart for 2006

CITL Professors **(numerical reported in %)**

Assessment		Knowledge	Comprehension	Application	Analyze	Synthesize	Evaluate
		Remember	Understand	Apply	Analyze	Evaluate	Create
Midterm	(1)	15%	30%	7.5%	7.5%	35%	None
	(2)	1	15	7	69	7	None
	(3)	5	35	45	10	5	None
	(4)	10	35	50	5	None	None
	(5)	None	10	50	40	None	None
	(6)	8	38	54	None	None	None
	(7)	30	70	none	none	none	none
Final	(1)	none	24%	none	12%	24%	46%
	(2)	2	14	8	76	None	None
	(3)	NR	NR	NR	NR	NR	NR
	(4)	5	30	55	10	None	None
	(5)	None	10	50	30	10	None
	(6)	4	35	61	None	None	None
	(7)	30	70	none	none	none	none
Performance Task 1	(1)	10%	10%	10%	20%	20%	30%
	(2)	None	None	10	30	30	30
	(3)	None	None	None	50	50	None
	(4)	None	None	None	None	100	None
	(5)	None	None	20	20	50	10
	(6)	None	None	None	None	50	50
	(7)	None	none	none	none	50	50

Assessment	Knowledge	Comprehension	Application	Analyze	Synthesize	Evaluate
	Remember	Understand	Apply	Analyze	Evaluate	Create
Homework (1)	**none**	**40%**	**5%**	**30%**	**10%**	**15%**
(2)	**10**	**20**	**30**	**40**	**none**	**none**
(5)	**None**	**30**	**50**	**20**	**None**	**None**
(6)	**None**	**40**	**30**	**30**	**None**	**None**
(7)	**X**	**X**	**none**	**none**	**none**	**none**
Lab Experiments (2)	**none**	**20**	**40**	**30**	**10**	**none**
<u>Miscellaneous:</u>						
Individual Assessments (4) (6)	**none**	**none**	**none**	**33**	**33**	**33**
Group Assess-Ments (4) (6)	**none**	**none**	**none**	**33**	**33**	**33**
Group Discussion (5)	**None**	**None**	**20**	**50**	**30**	**None**
Round Table Discussion (5)	**None**	**None**	**20**	**50**	**30**	**None**
Oral Presentation (5)	**none**	**none**	**none**	**10**	**40**	**50**

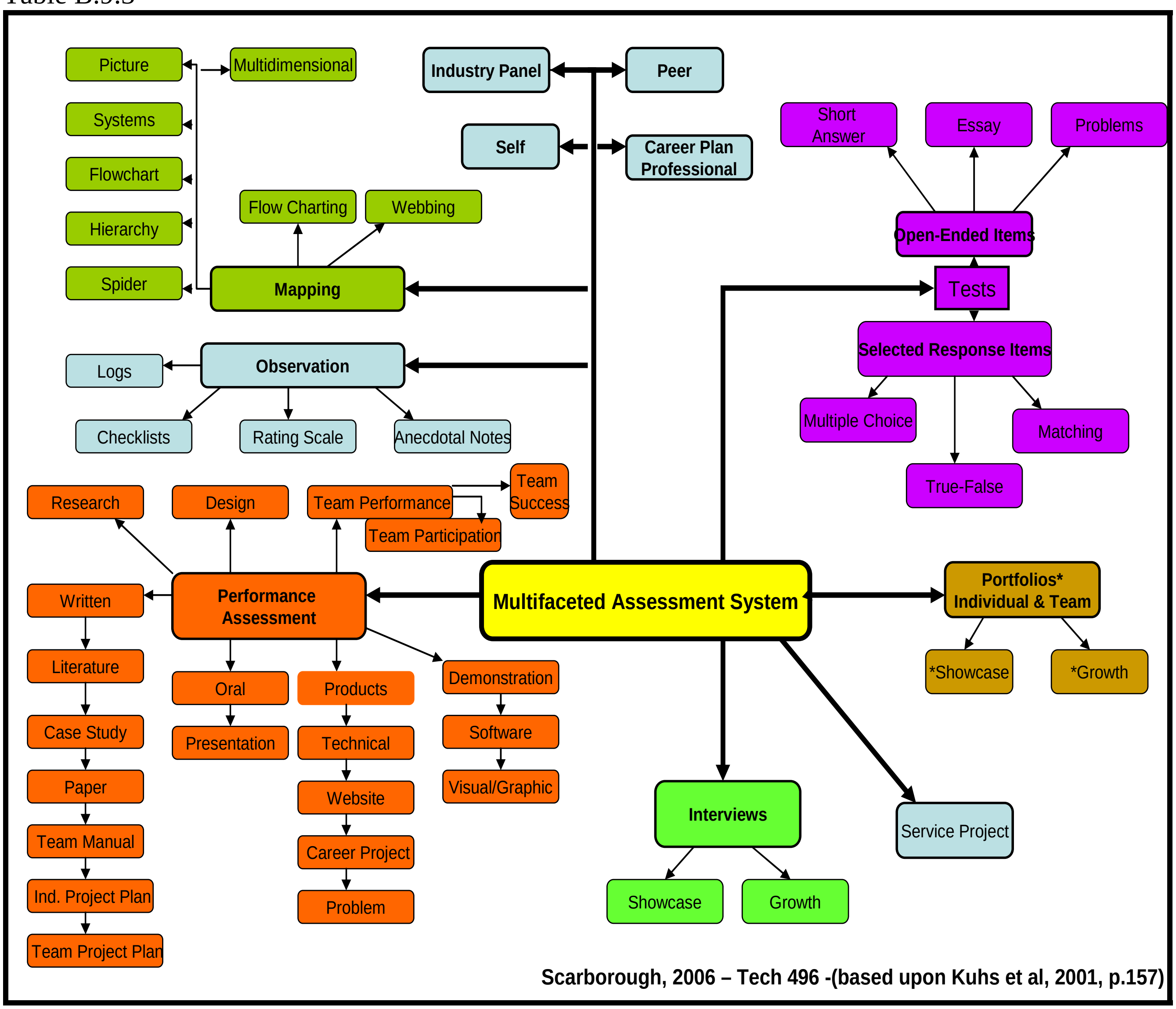
During the 2006 experimental research semester, the professors used their new assessments to measure and provide evidence of student learning. After the midterm and final exams, they performed an analysis, as they did initially with the 2005 tests. This process moved them into further use of test analysis for diagnostic purposes to improve tests, instruction, and student learning. They also performed their first diagnostic review of the performance tasks, qualitatively, to determine strengths and needed improvements. Both of these processes will assist them in improving the assessments for the next course offering and further instill the best practices of test analysis for diagnostic purposes, the use of performance tasks/rubrics, and the analysis of the performance tasks/rubrics (also for diagnostic purposes) to improve assessment and instruction, and ultimately student learning.

This was a successful program component, and there was significant gain in knowledge and skills by all professors. See following sections:

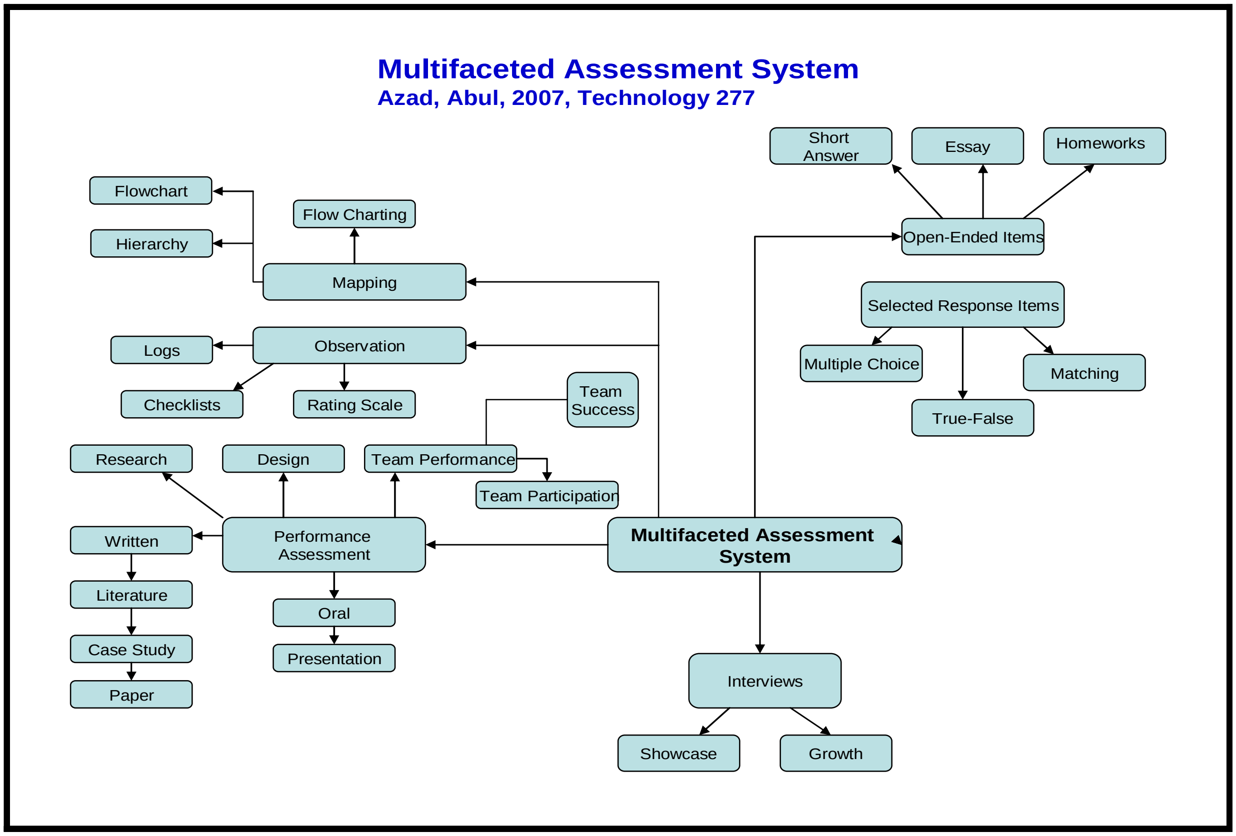
- **Bloom's Analysis of Assessment Plans (copied into this document above)**
- **Final Report: Discussion of Professors' 2005 and 2006 Test Analyses and Tests (See B.9.a)**
- **Professor Diagnostic Analyses of their 2006 midterm and final exams. (See B.9.d)**
- **Professors' Multifaceted Assessment Plan Graphics (samples are copied into this document below) The Maps are based upon Kuhs et al. (2001) and Scarborough (2006) modifications for the Faculty Development Program (Tables B.9.3). The Program sample is below, followed by each of the Faculty participants (Tables B.9.4-9).**

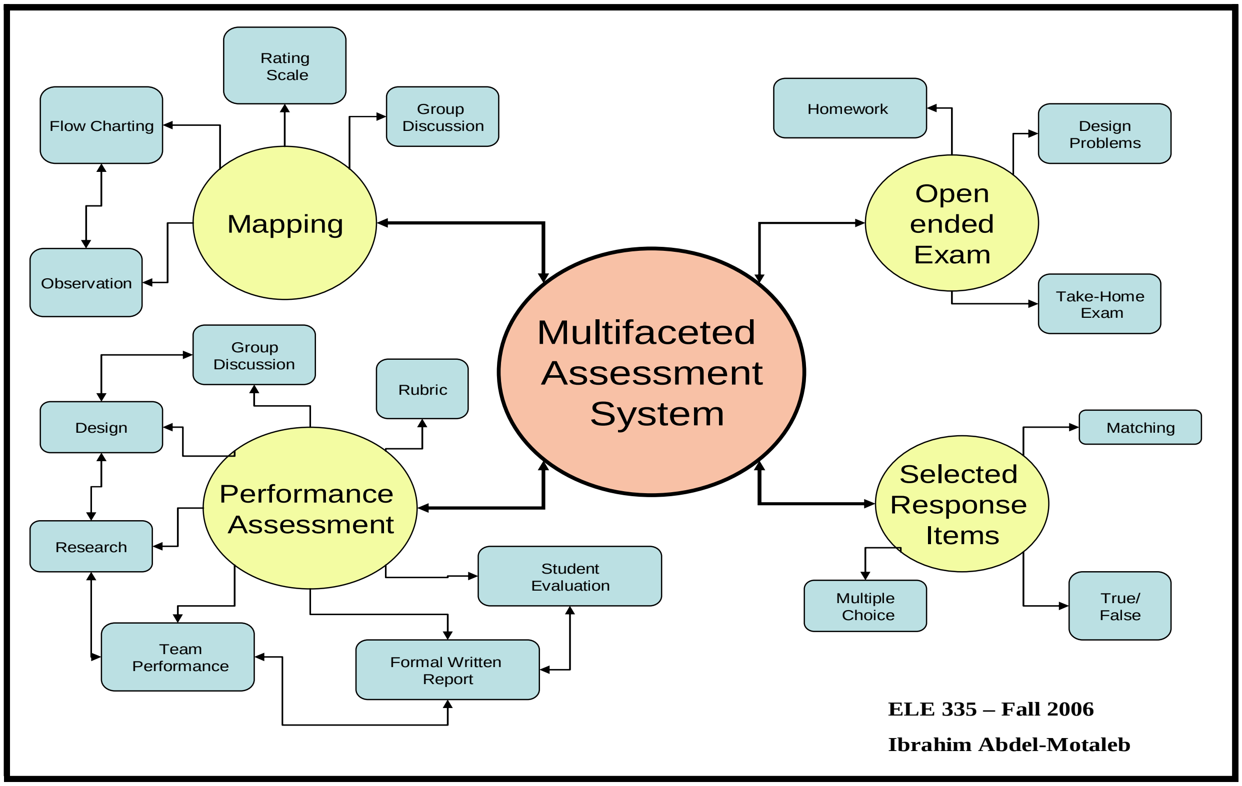
Note: We were going to explore studying the correlations between the Objective Tests and the Performance Assessments. We did compute the correlations, but the research considerations were too complex to consider. However refer to B.9.c for the discussion.

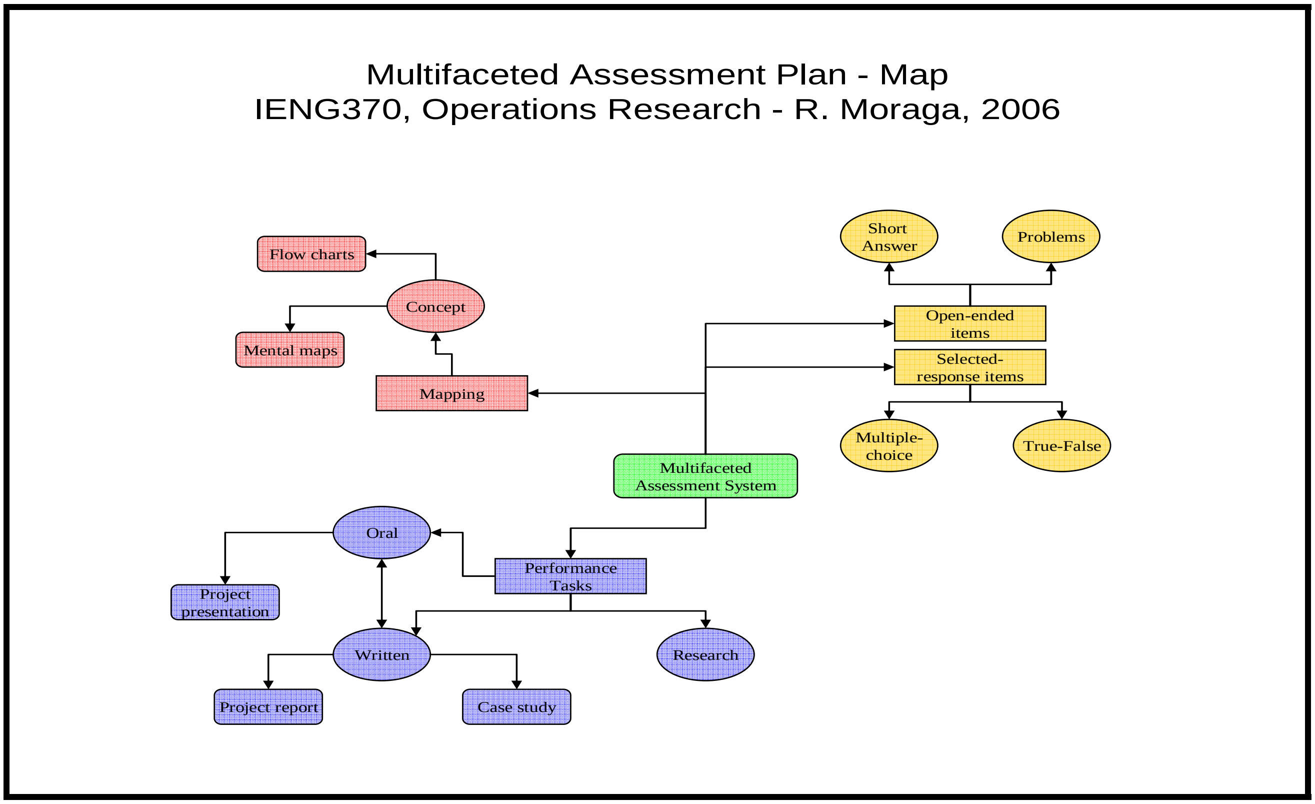
Table B.9.3



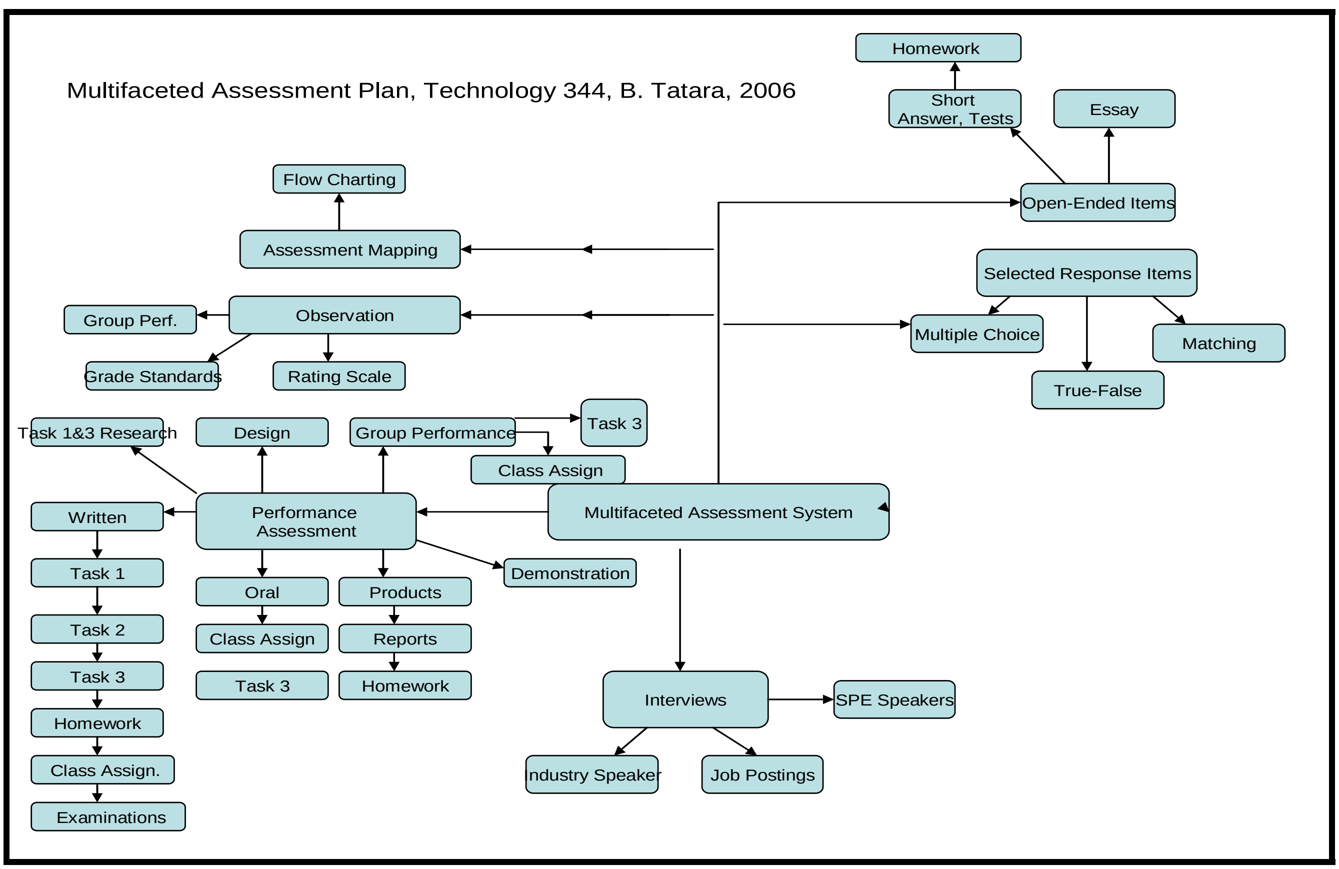
CEET Professors 2006 Multifaceted Assessment Plans
Maps based on Scarborough, 2006 and Kuhs et al, 2001



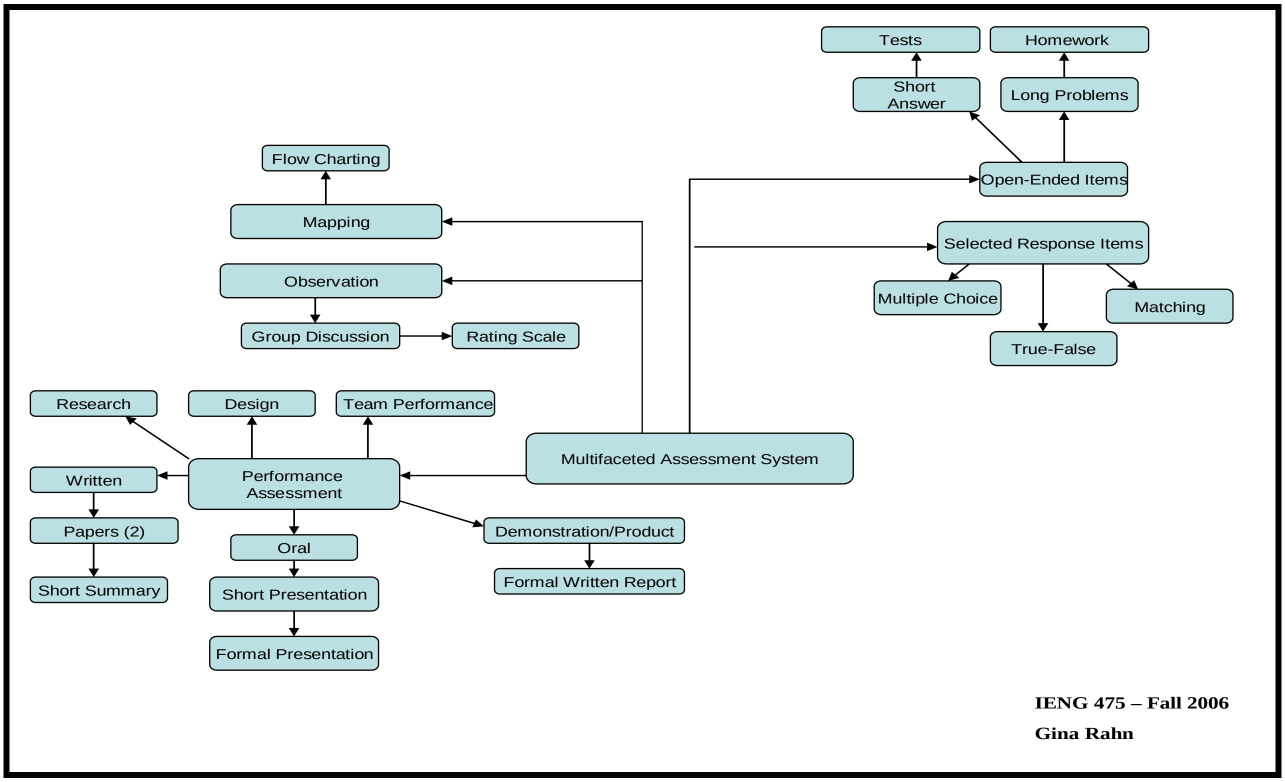


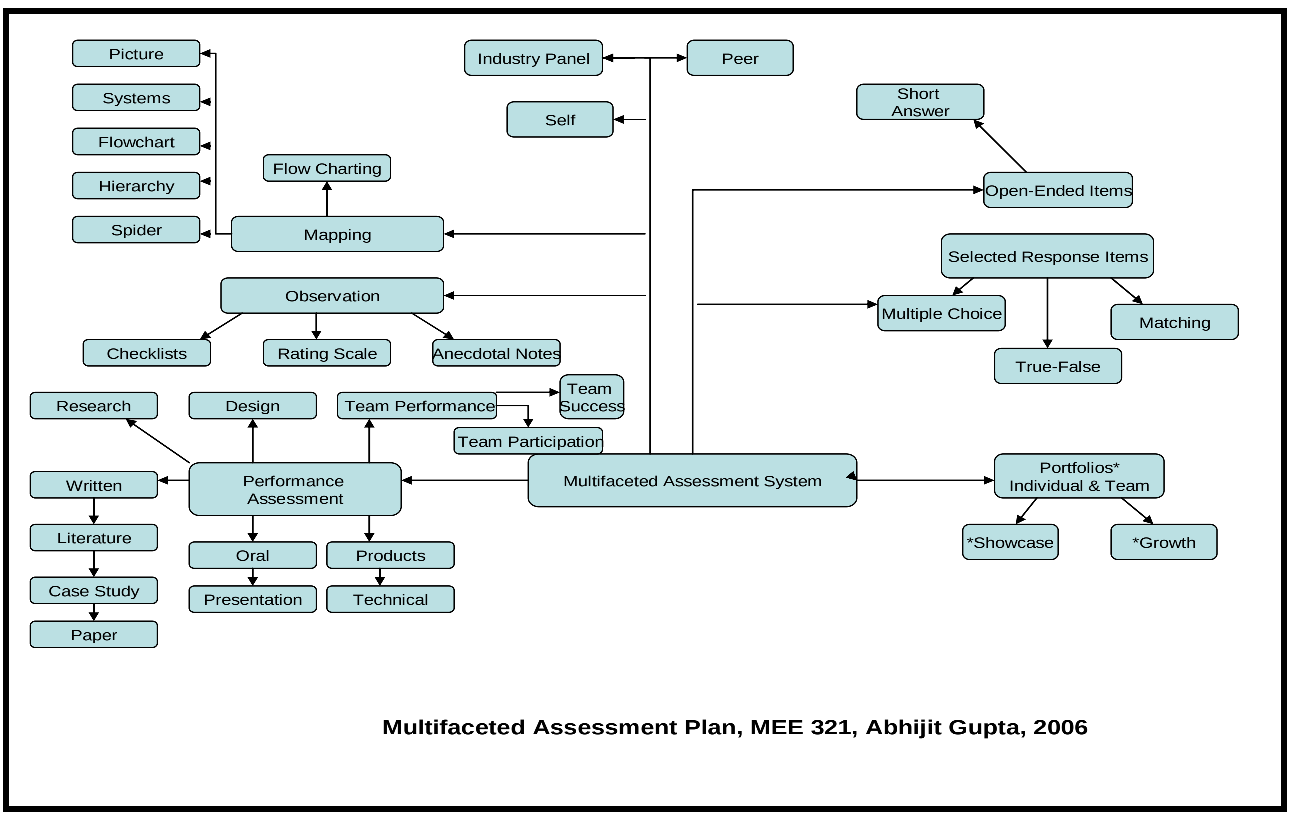


Multifaceted Assessment Plan - Map
IENG370, Operations Research - R. Moraga, 2006
Flow charts
Concept
Mental maps
Mapping
Short Answer
Problems
Open-ended items
Selected-response items
Multiple-choice
True-False
Multifaceted Assessment System
Oral
Performance Tasks
Project presentation
Written
Research
Project report
Case study



Multifaceted Assessment Plan, Technology 344, B. Tatara, 2006
Homework
Short Answer, Tests
Essay
Flow Charting
Open-Ended Items
Assessment Mapping
Selected Response Items
Group Perf.
Observation
Multiple Choice
Matching
Grade Standards
Rating Scale
True-False
Task 1&3 Research
Design
Group Performance
Task 3
Class Assign
Written
Performance Assessment
Multifaceted Assessment System
Task 1
Oral
Products
Demonstration
Task 2
Class Assign
Reports
Task 3
Task 3
Homework
Homework
Interviews
SPE Speakers
Class Assign.
Industry Speaker
Job Postings
Examinations


Tests
Homework
Short Answer
Long Problems
Flow Charting
Open-Ended Items
Mapping
Selected Response Items
Observation
Multiple Choice
Matching
Group Discussion
Rating Scale
True-False
Research
Design
Team Performance
Multifaceted Assessment System
Written
Performance Assessment
Papers (2)
Demonstration/Product
Oral
Short Summary
Short Presentation
Formal Written Report
Formal Presentation
IENG 475 – Fall 2006
Gina Rahn


Picture
Industry Panel
Peer
Systems
Self
Short Answer
Flowchart
Flow Charting
Hierarchy
Open-Ended Items
Spider
Mapping
Selected Response Items
Observation
Multiple Choice
Matching
Checklists
Rating Scale
Anecdotal Notes
True-False
Team Success
Research
Design
Team Performance
Team Participation
Written
Performance Assessment
Multifaceted Assessment System
Portfolios* Individual & Team
Literature
Oral
Products
*Showcase
*Growth
Case Study
Presentation
Technical
Paper
Multifaceted Assessment Plan, MEE 321, Abhijit Gupta, 2006