

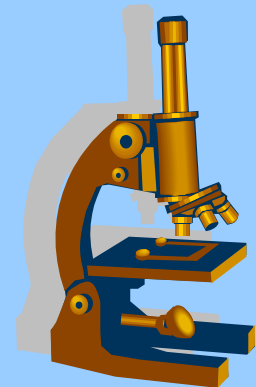
**IR and the Challenges of Learning Outcomes
Assessment: Case Study of an Institutional
Research Office's Collaboration with the Writing
Faculty to Assess the College's Writing Program**

**Gail Wisan, Ph.D.
Director
Office of Institutional Research, Goucher College**

**Pallabi Roy, Ph.D.
Assistant Director of Institutional Research for
Assessment and Research, Goucher College**

**Presented at the Annual Conference of
The Association for Institutional Research**

**Seattle, Washington
May 28, 2008**



Learning Outcomes Assessment: Context

National:

- At the National Symposium on Student Success, Secretary of Education Margaret Spellings called on colleges to measure and provide evidence of student learning.
- The outcomes assessment has two purposes*
 - "Accountability"
 - "Effectiveness": Important Indicator of:
Student Learning Outcomes: Writing Ability

* 'Characteristics of Excellence' – The Middle States Commission on Higher Education

Learning Outcomes Assessment: Context

- Local:
 - Small liberal arts college
 - Academic Dean placed assessment function in IR
 - Goucher's Student Learning Outcomes Assessment (SLOA) plan cycle was developed.
 - SLOA included writing program in first 2 years.

Writing Outcomes Assessment: Context



■ Writing Program Assessment:

❖ Faculty had concerns about

- the value of learning outcomes assessment
- the time commitment

Writing Outcomes Assessment: Context

- Writing Program Places Each Incoming Student in Core Writing Class:
 - All entering first year students are required to take department developed web based writing prompt.
 - Required writing placement tests serve as a writing pretest.
 - Students move through the required writing courses (Academic Writing) until they achieve writing proficiency

Writing Outcomes Assessment: Context

- Writing Program Assessment:
 - Required writing placement tests served as a writing pretest for IR writing assessment.
 - Placement in writing class serves as a “grade” on the placement test
 - Writing faculty did Not use a rubric for student placement
 - Writing faculty did Not record detail on each placed student’s performance on the writing placement essay —only placement in English 103, 104 or 105.

Writing Program Assessment: Context

- This session reports on a case study of how an IR Office developed a collaborative process with the English writing faculty to assess students' writing skills in the college's core required writing courses:

English 104: placement of average writers

English 103: placement of weak writers

English 105: placement of strong writers

English 106: placement of transfer students and students without proficiency in Eng. 105

The Writing Program Assessment Process: The Population and Sample

- **Two Phases of the Assessment Process:**
2006-2007 and 2007-2008


Population:

Students in Core Proficiency Writing Courses--

English 103, 104, 105, 106 writing classes


- **Sample:** papers were randomly selected from each course. (Details later)

The Writing Program Assessment Process: The Hypotheses



- Research Question: Is the Writing Program Placement Process working successfully?
- Research Hypothesis: English 105 students will perform at a higher level than English 104 students on all rubric dimensions.
- Research Hypothesis: English 104 students will perform at a higher level than English 103 students on all rubric dimensions.
- Null Hypothesis: All classes will perform the same on all writing dimensions.

The Writing Program Assessment Process: The Rubric



- IR Meets with Writing Faculty
- The Rubric: Defining and Measuring Student Learning Outcomes for the Writing Program
- Writing faculty were not satisfied with writing rubric but thought they were “stuck” with IR provided rubric
- IR told them they could modify the writing rubric.
- Subgroup of English faculty met to revise writing rubric.

Dimensional Rubric for Writing Assessment

Element	Inadequate (1)	Very Weak (2)	Adequate (3)	Competent (4)	Strong (5)	<i>Outstanding</i> (6)	Comments
1. Control A <i>strong</i> paper has an immediate direct or implied sense of purpose that controls the content and shape of the paper. The purpose is sufficiently concrete and appropriate in scope for the paper's intended audience. The content sustains a logical organization, demonstrated through paragraph structure and the overall shape of the essay.							

Dimensional Rubric for Writing Assessment

Element	Inadequate (1)	Very Weak (2)	Adequate (3)	Competent (4)	Strong (5)	Outstanding (6)	Comments
<p>2. Thought</p> <p>A <i>strong</i> paper presents arguments, ideas and information that are meaningful and substantial. For academic discourse, it has a thesis that is complex, significant and specific, synthesizing and interpreting information from sources rather than merely echoing them. For other audiences, it demonstrates a creative and distinct approach to the material. It may render an unfamiliar topic cogent and memorable, or a familiar topic fresh and incisive.</p>							

Dimensional Rubric for Writing Assessment

Element	Inadequate (1)	Very Weak (2)	Adequate (3)	Competent (4)	Strong (5)	<i>Outstanding</i> (6)	<i>Comments</i>
3. Development A <i>strong</i> paper is developed logically, is cognizant of multiple points of view, and uses credible evidence and well chosen details to support its claims. Outside sources, if used, are evaluated carefully, presented in context, integrated smoothly and persuasively, and cited correctly.							

Dimensional Rubric for Writing Assessment

Element	Inadequate (1)	Very Weak (2)	Adequate (3)	Competent (4)	Strong (5)	<i>Outstanding</i> (6)	<i>Comments</i>
4. Language A <i>strong</i> paper is clear, concise, and pithy. There is a voice present, and that voice socializes with the reader, engaging and sustaining interest through artful construction and tension in language or argument. Sentences are varied, uncluttered and efficient. Standard usage is employed, and errors either are not present or do not interfere with meaning.							

Dimensional Rubric for Writing Assessment

Element	Inadequate (1)	Very Weak (2)	Adequate (3)	Competent (4)	Strong (5)	<i>Outstanding</i> (6)	<i>Comments</i>
5. Overall A <i>strong</i> paper presents and executes a careful plan for its material. Arguments are compelling, explanations are nuanced and lucid, descriptions are vivid, and language is exact. The reader is engaged in the text from start to finish and is left with a sense of fulfillment or artful resonance. This paper is a worthwhile experience for the reader, who has gained unexpected knowledge, insight, or perspective from it.							

Details of Writing Assessment: Methodology

- In January 2007, 164 papers from four different fall 2006 courses were evaluated by the English writing faculty. The break down by course is as follows:
 - ENG 103 [The College Essay]: 26 [74% of the sample of 35]
 - ENG 104[Academic Writing I]: 89 [71% of the sample of 125]
 - ENG 105[Academic Writing II]: 34 [97% of sample of 35]
 - ENG 106[Academic Writing III]: 15 [43% of sample of 35]
 - **Total of all classes: 164 [71% of 230]**


Details of Writing Assessment: Methodology

- In May 2007, 164 papers from 4 different spring 2007 courses were evaluated by the English writing faculty. The break down by course is as follows:
 - ENG 104: 25 [61% of all 41 enrolled]
 - ENG 105: 134 [all 84 students assessed in Fall 2006 plus 50 more students [48% of sample of remaining 104]]
 - ENG 106: 5 [21% of all 24 enrolled]
 - **Total of all classes: 164 [65% of 253 sampled]**

First Phase of Writing Assessment: Methodology

- ❖ Each paper (both January and May assessments) was graded on a six point scale using the rubric with five dimensions: control, thought, development, language, and overall.
- ❖ IR retrieved students' final papers submitted in BlackBoard and erased any identifiers e.g. names, course number etc.
- ❖ Each paper was graded anonymously by 2 different writing faculty.
- ❖ In late January 2007, the writing faculty assessed the 164 fall writing papers during one long assessment day.
- ❖ In late May, the writing faculty scored 164 spring 2007 writing papers during two shorter assessment days.

Results – January 2007 Writing Assessment of Fall 2006 Freshmen



- Anonymous grading indicated that students are being placed correctly in their writing classes.
- English 105 students ranked higher than English 104 students on all rubric dimensions
- English 104 students ranked higher than English 103 students on all writing dimensions
- ENG 106 tended to rank between ENG 104 and ENG 105 students. [English 106 is mainly for transfer students.]

Results by Rubric Dimension – Fall 2006 Assessment

Control

- **English 105** and **English 106** students' papers averaged **competent** (mean = 4.1 for ENG 105 and 3.9 for Eng 106).
- **English 104 and 103** are rated between **adequate and competent** [3.7 for English 104 and 3.2 for English 103]

Thought

- **English 103**: Between very weak and adequate (2.7).
- **English 104**: Adequate (mean = 3.4).
- **English 105**: Competent (mean = 4)
- **English 106**: Somewhat below competent (mean = 3.7).

Development

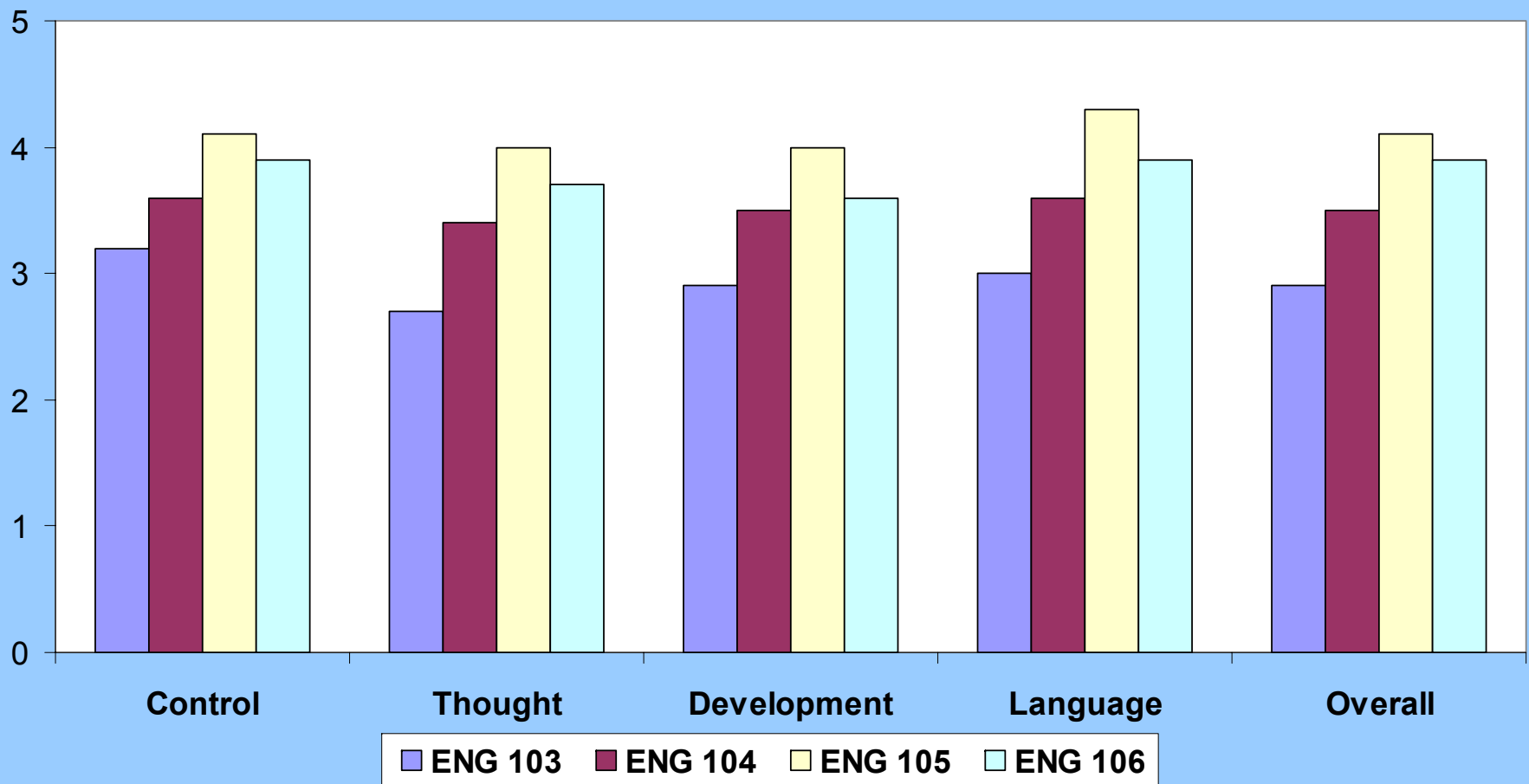
- **English 103** students averaged slightly below competent (mean = 2.9).
- **English 105** students averaged the score of competent.
- **English 106** students are somewhat lower, followed by 104 students.

Language

- **English 105 and 106** students averaged above and below competent respectively.
- **English 103** students average levels were adequate, whereas **English 104** levels were between adequate and competent.

Overall - The **105** (mean = 4.1) and **106** (mean = 3.9) papers received competent overall assessment while 103 scored about adequate (mean = 2.9) and **English 104** in between adequate and competent (mean = 3.5).

Fall '06 Writing Program Assessment: Learning Outcomes Mean Scores by Course



Results by Rubric Dimension – Spring 2007 Writing Program Assessment

❑ Control

- ❑ ENG 104 papers are 0.1 less than adequate
- ❑ English 105 and ENG 106 are rated between adequate and competent with ENG 105 scoring more towards competent.

❑ Thought

- ❑ English 104 students are between very weak and adequate in the direction of adequate.
- ❑ English 105 students are better than adequate
- ❑ English 106 students are just adequate.

❑ Development

- ❑ English 104 students have an average score of somewhat less than adequate.
- ❑ English 105 scores are higher, followed by 106 students who were assessed at a level close to competent.

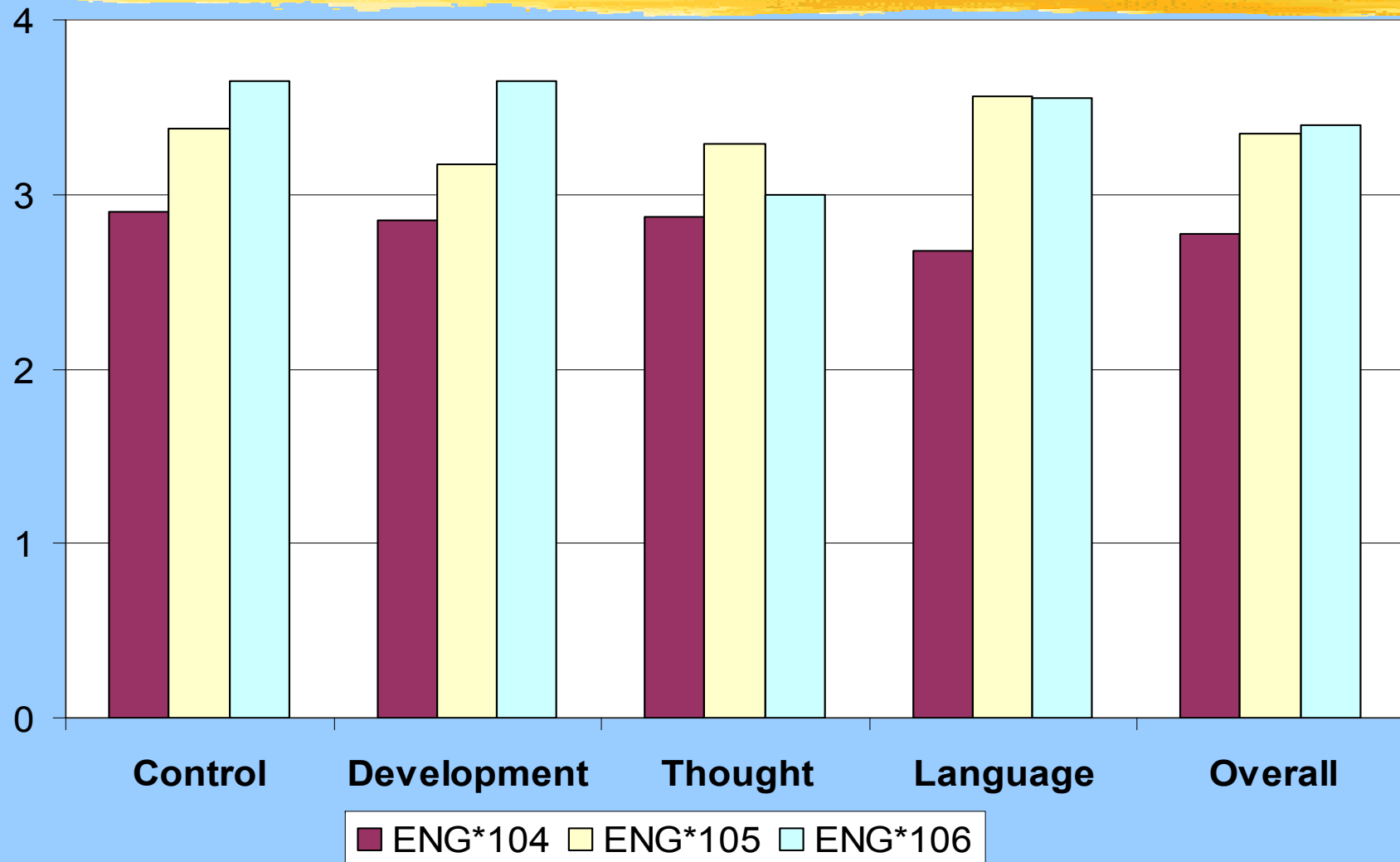
❑ Language

- ❑ English 105 and 106 students performed at the same level between adequate and competent.
- ❑ English 104 students average just below adequate.

❑ Overall

- ❑ The 105 and 106 papers obtained the overall assessment in the middle of adequate and competent,
- ❑ The 104 papers received close to adequate.

Spring '07 Writing Program Assessment: Learning Outcomes Mean Scores by Course



Results: Consistency of Evaluators

- **For fall 2006 writing papers**, a difference between evaluator #1 and evaluator #2 of 0.5 or more (in average score) occurred only in the Development dimension.
 - **Overall, evaluators graded consistently** in terms of the mean scores of students on rubric dimensions.
- **For spring 2007 writing assessment**, there was somewhat less consistency English 106 papers ratings:
 - Ratings of English 106 papers had differences of 0.5 or more between evaluators on all dimensions except control with the highest difference being a full point, 1.0, in the overall dimension.
 - English 104 and English 105 student scores were fairly consistent between evaluators #1 and #2.

Fall '06 Writing Program Assessment Statistics by Evaluator

		Control				Thought				Development				Language				Overall			
	Evaluat or	Min	Ma x	Me an	Med ian	Mi n	M ax	Me an	Medi an	Mi n	M ax	Me an	Medi an	Mi n	M ax	Me an	Medi an	Mi n	M ax	Me an	Med ian
E N G 1 0 3	#1	2.0	5.0	3.2	3.0	1.0	6.0	2.8	2.5	2.0	5.0	3.1	3.0	2.0	6.0	3.2	3.0	2.0	6.0	3.0	3.0
	#2	1.0	5.0	3.2	3.0	1.0	4.5	2.6	3.0	1.0	5.0	2.7	2.5	1.0	5.0	2.9	3.0	1.0	5.0	2.8	3.0
E N G 1 0 4	#1	1.0	6.0	3.7	4.0	2.0	6.0	3.5	3.0	1.0	6.0	3.5	4.0	1.0	6.0	3.7	4.0	1.0	6.0	3.6	4.0
	#2	1.0	6.0	3.6	4.0	1.0	6.0	3.3	3.0	1.0	6.0	3.4	3.0	1.0	6.0	3.5	3.0	1.0	6.0	3.4	3.0
E N G 1 0 5	#1	2.0	6.0	4.0	4.0	2.0	6.0	4.0	4.0	2.0	6.0	4.0	4.0	3.0	6.0	4.3	4.0	2.0	6.0	4.0	4.0
	#2	2.0	6.0	4.2	4.0	1.0	6.0	4.0	4.0	1.0	6.0	3.9	4.0	2.0	6.0	4.3	5.0	1.5	6.0	4.1	4.0
E N G 1 0 6	#1	2.0	5.0	4.0	4.0	1.0	5.0	3.7	4.0	2.0	5.0	3.9	4.0	3.0	5.0	3.9	4.0	2.0	5.0	4.0	4.0
	#2	2.0	6.0	3.7	4.0	2.0	6.0	3.7	4.0	1.0	6.0	3.2	3.0	2.0	6.0	3.9	4.0	2.0	6.0	3.7	4.0

Discussion – May 2007 Writing Assessment of Students in Spring 2007 Courses

- Overall rubric dimension means for students in a course (e.g., Eng.104) were **lower** in May 2007
- Faculty disappointed
- English 105 performed the best, except for Development where it trails behind English 106 by a margin of 0.5 (In the Fall, English 105 were clearly rated higher than English 106)
- English 104 stands lower than English 105 and 106 in all dimensions

Discussion (Cont.) – May 2007 Writing Assessment of Students in Spring 2007 Courses

- Spring 2007 English Writing Students not exactly the same population as fall 2006
- Highest achieving students already achieved writing proficiency and are not in the spring 2007 core writing classes
- Best writing students were in English 105 in fall 2006. Spring 2007 English 105 students are English 104 students moving up to the next level.

Comparison of Writing Skills/Learning Outcomes by Gender

- Of the 164 fall 2006 assessed students, 32% were male.
 - English 103 and English 105 male students mean rubric scores were higher than female students mean scores except on the *language dimension for ENG 103*.
 - For English 104 and 106, the gender trend seemed to move in the opposite direction; females often scored higher than males on writing skills/learning outcome rubric dimensions.

- For spring 2007, the gender distribution of the assessed students shows 36% male, a 4% increase from fall '06.
 - ENG 104 and ENG 105 male and female students scored the same in most dimensions except a highest difference of 0.2 occurring in 'thought'.
 - ENG 106 sample has no male students.

Comparison of Writing Skills by Gender

T-test comparison of mean scores by gender

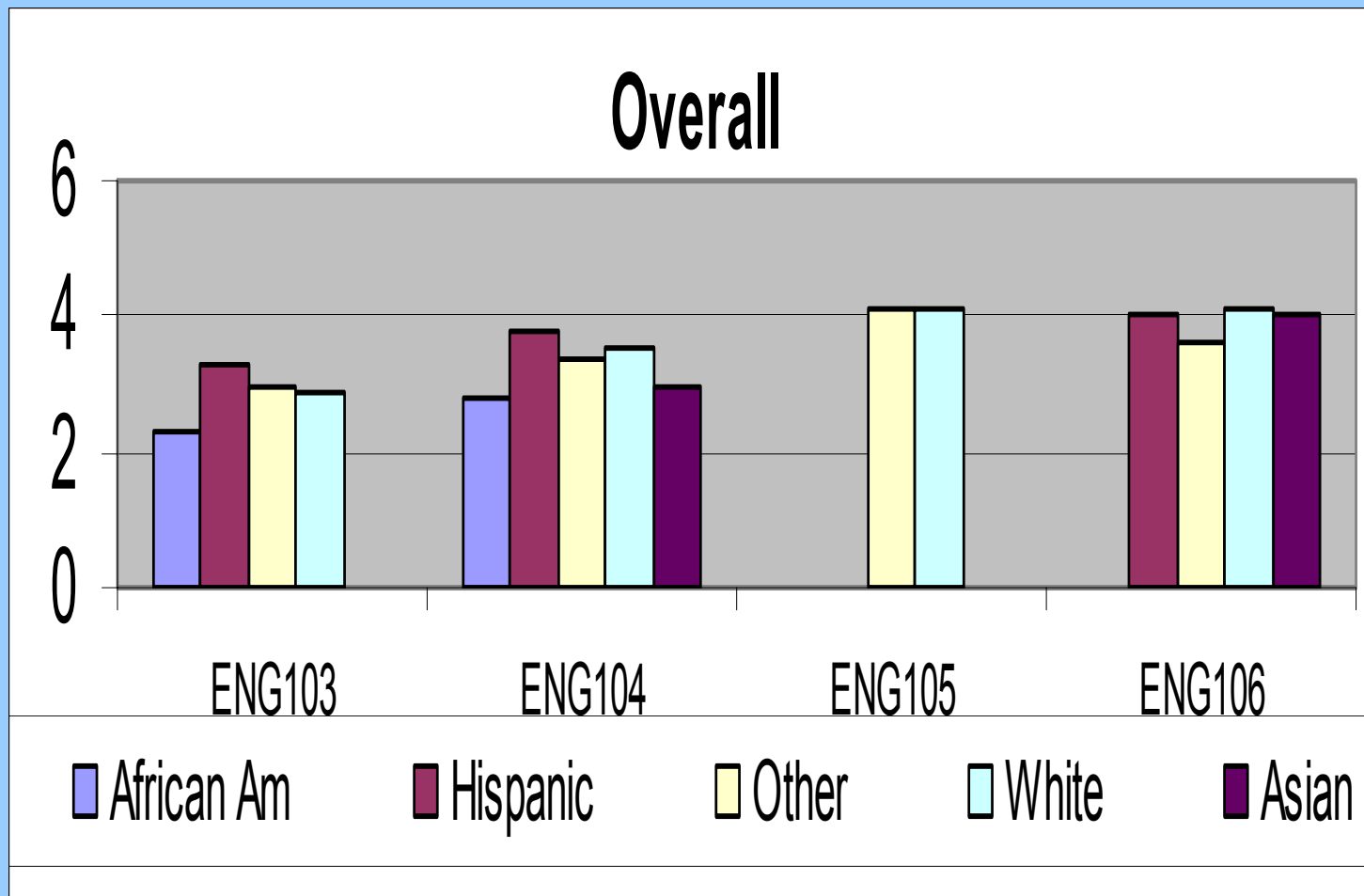
- Used t-tests to evaluate whether the apparent differences constituted real differences as opposed to minor chance variations.
- Fall 2006:
 - English 104 female mean scores were found to be higher on every ranking dimension.
 - Gender differences in mean scores reached the level of statistical significance ($p \leq .05$) on all rubric dimensions with the exception of *language*.
 - Apparent higher scores among males in English 103 and English 105 did not reach the level of statistical significance. In English 105, males mean score on *thought* was almost significant, barely missing the .05 probability level.
- Spring 2007:
 - No differences in mean scores by gender were found to be statistically significant and higher scores on dimensions bounced back and forth between males and females.

Comparison of Writing Skills/Learning Outcomes by Ethnicity/Race

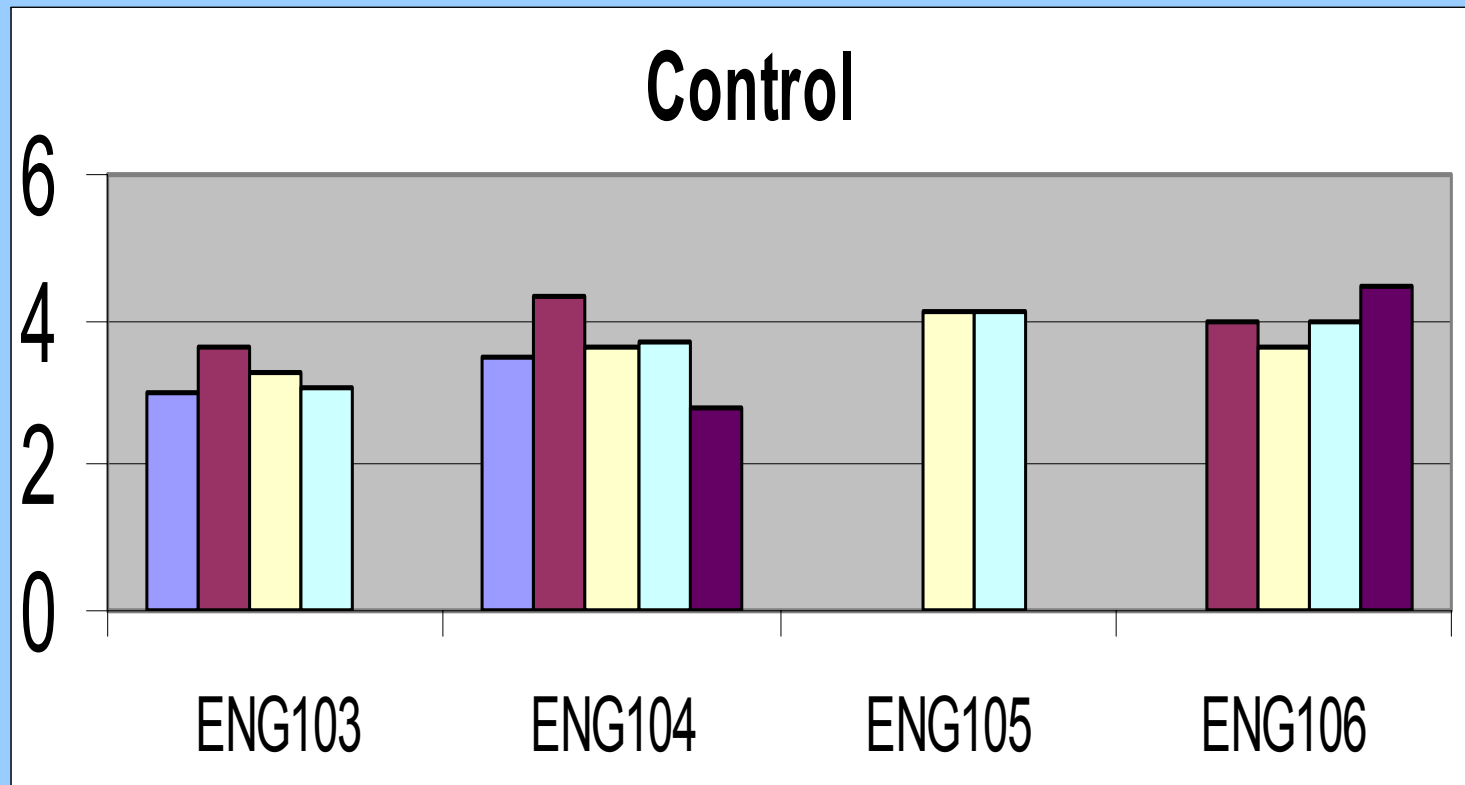
Fall 2006:

- 73% White
- 3% African American
- 4% Hispanic
- 3% Asian
- 17% Unknown Race
- No race comparison is available within the English 105 course since no African American, Asian, or Hispanic students were in the English 105 sample.
- The race chart shows a comparison of scores for four main ethnicities, White, African American, Hispanic, and Asian.

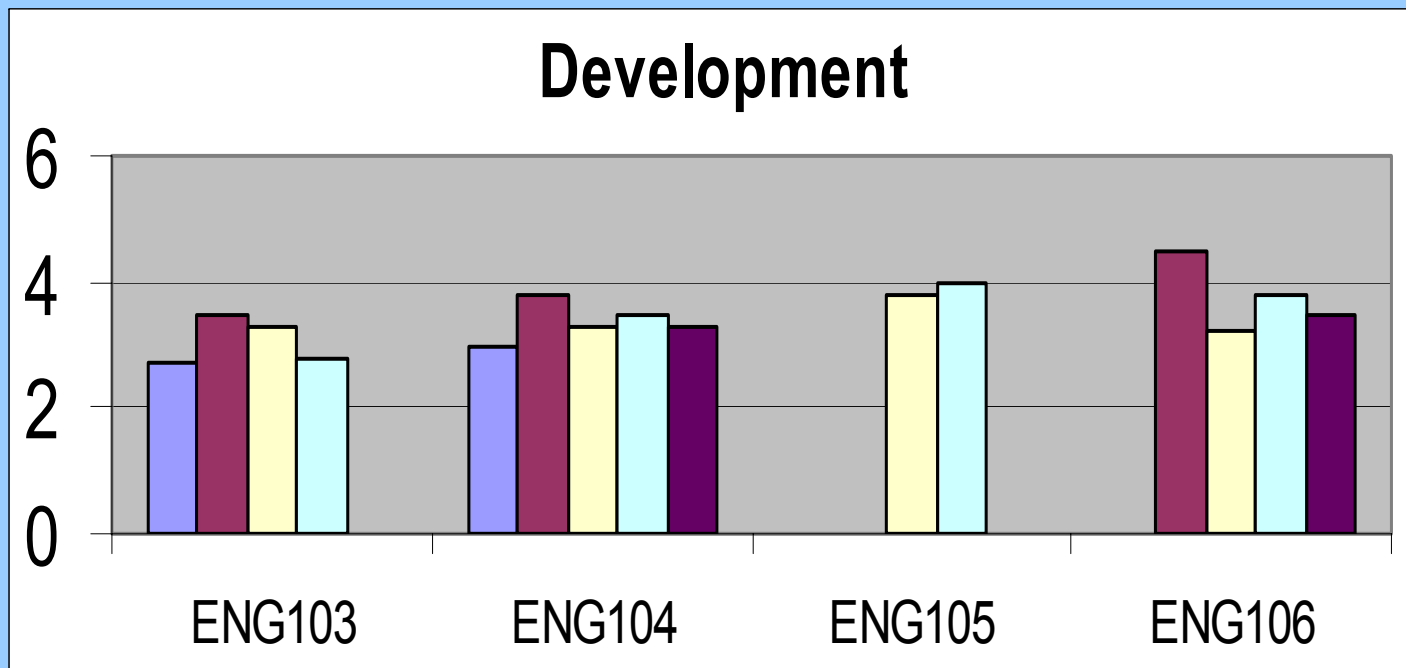
Fall 2006 Writing Program Rubric Dimensions: Overall by Course and Race



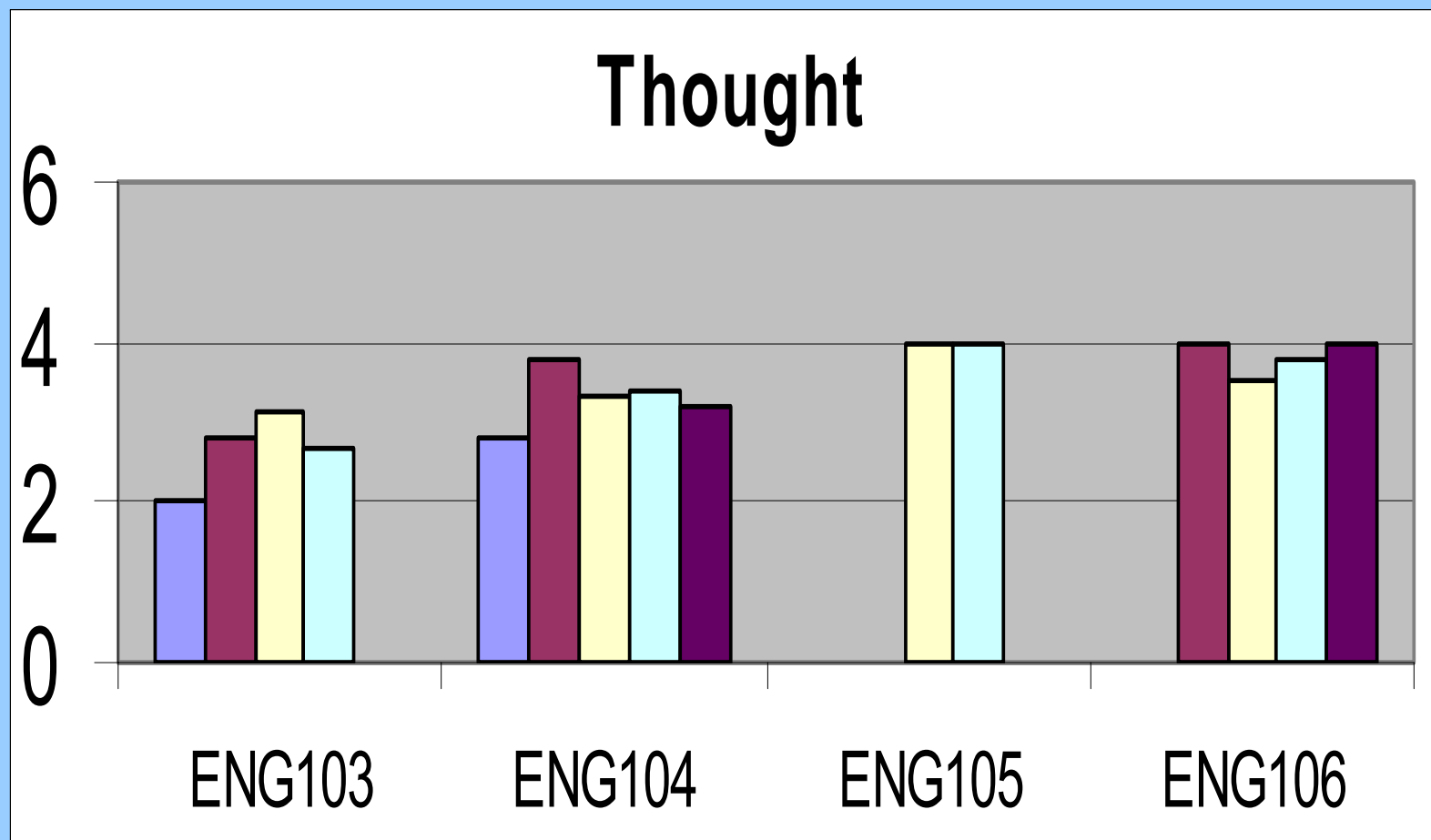
Fall 2006 Writing Program Rubric Dimensions: Control by Course and Race



Fall 2006 Writing Program Rubric Dimensions: Development by Course and Race

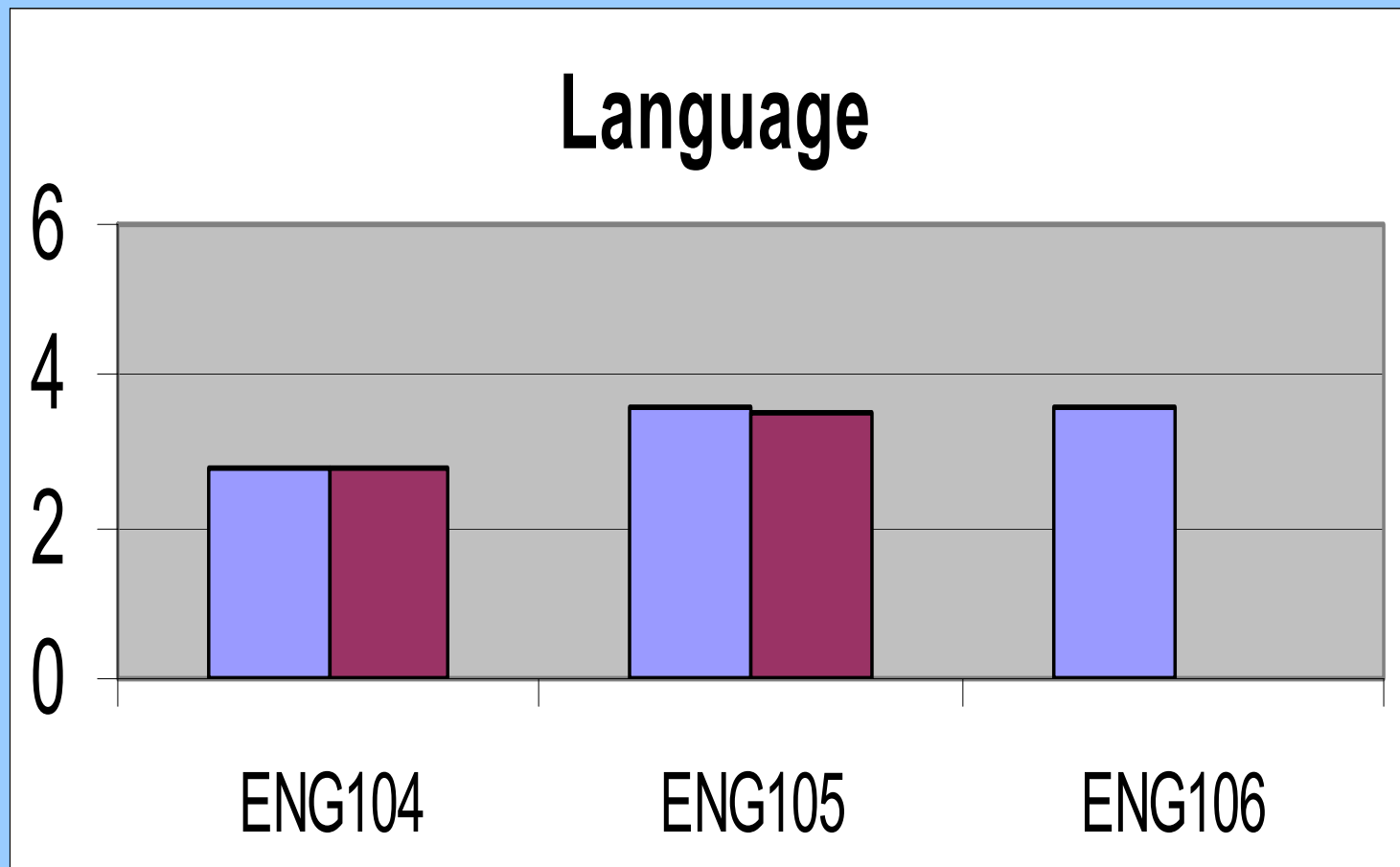


Fall 2006 Writing Program Rubric Dimensions: Thought by Course and Race



Fall 2006 Writing Program Rubric

Dimensions: Language by Course and Race




Comparison of Writing Skills/Learning Outcomes by Ethnicity/Race

Spring 2007:

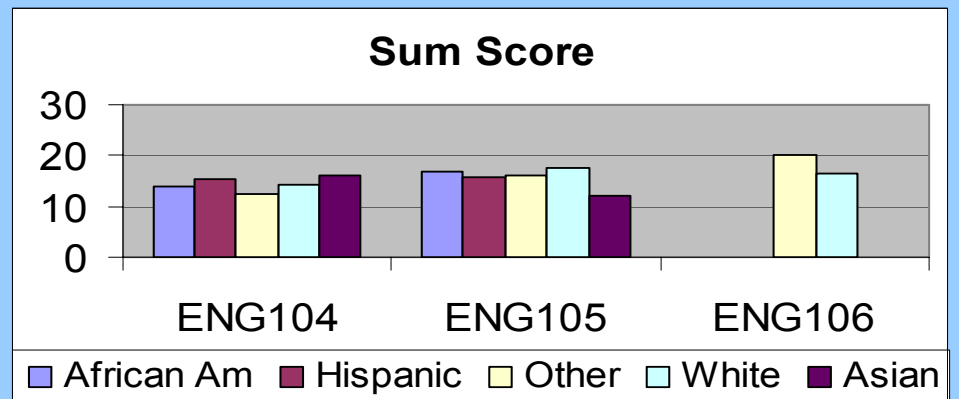
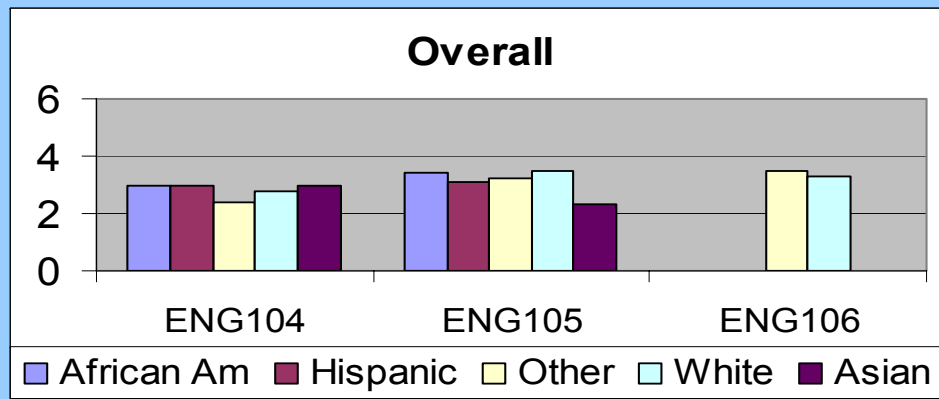
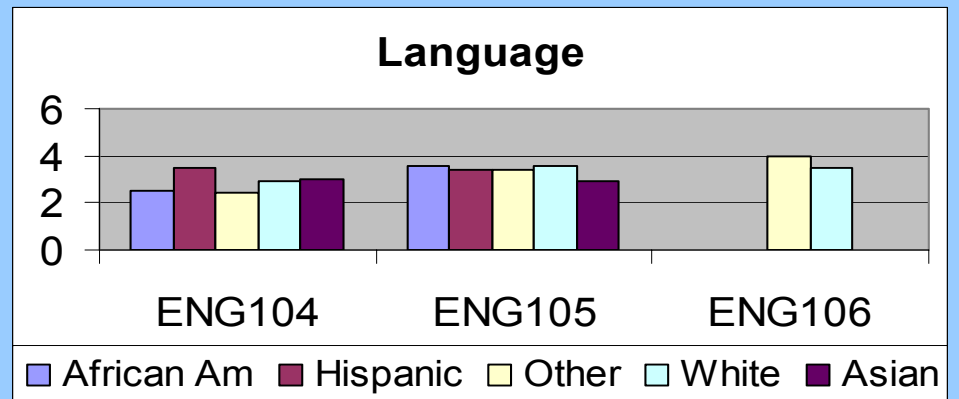
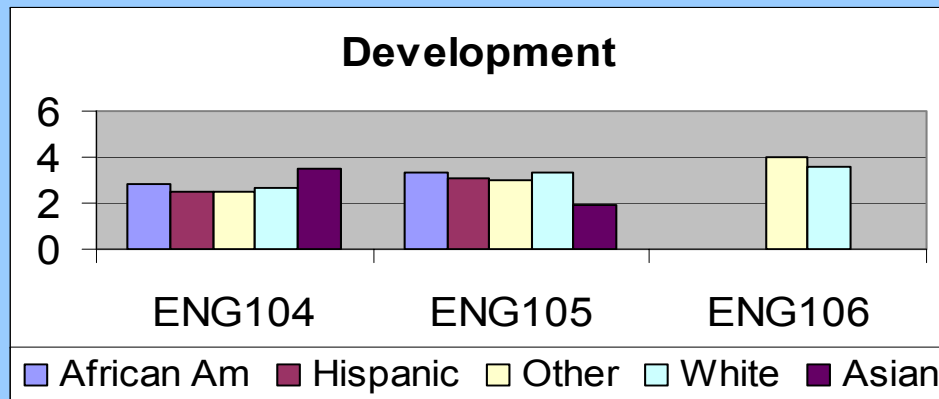
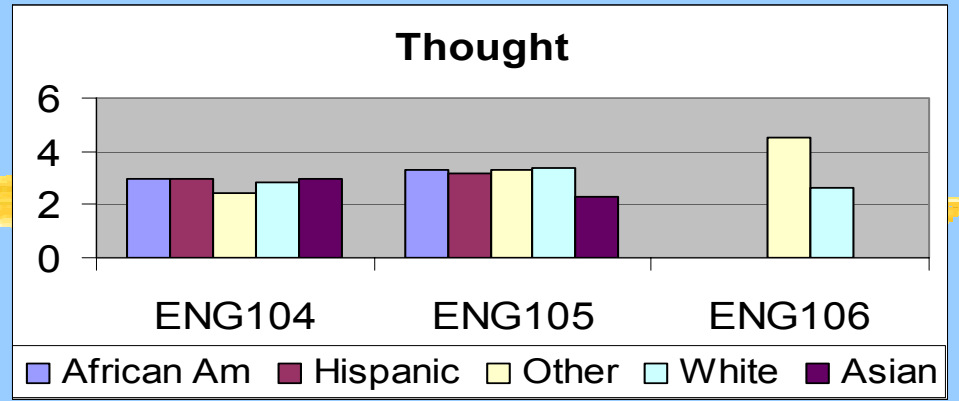
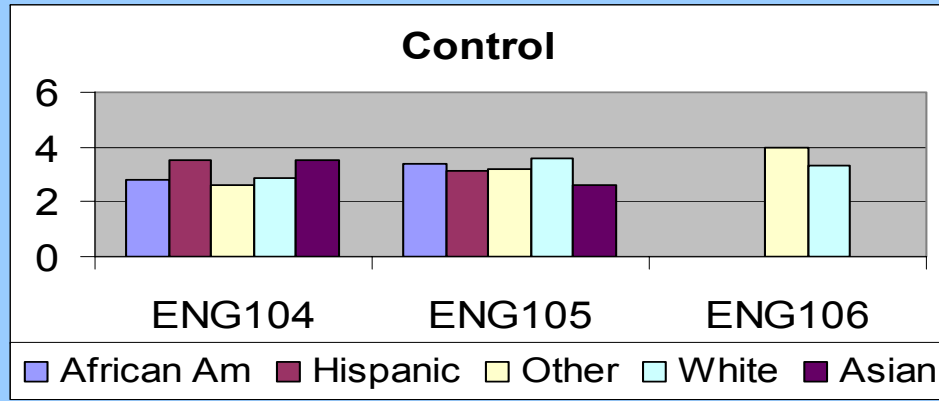
- 70% White
- 7% African American
- 4% Hispanic
- 2% Asian
- 17% Unknown Race
- No statistically significant differences were found
- **Hispanics again displayed a pattern of superior ratings on rubric dimensions**

Serendipity: Interesting AND Usable Research Result



- Hispanic Students performed at higher levels than other students in English 103 and English 104.
- Placement is NOT completed anonymously--- Hispanic names are seen by faculty doing placement.
- Question: Are Hispanic students being placed low because of low evaluations affected by their names?
- Recommend: Reinstigate anonymous placement

Spring '07 Writing Program Learning Outcomes by Course and Race



Results Fall 2006: Correlation of Rubric Score Sums with SAT Scores and High School GPA

Variable	Mean	Minimum	Maximum	Pearson Correlation of Writing Assessment Sum of Scores with SAT Scores
Sum Score	17.86	7.5	27.5	
SAT Essay	9.03	6	12	0.28232 **
SAT Verbal	632.47	390	800	0.44324***
SAT Writing	628.12	370	800	0.46096***
High School GPA	3.23	2.15	4	0.31279***

* significant at $\leq .05$

**significant at $\leq .01$

***sianificant at $\leq .001$

Results Fall 2007: Correlation of Rubric Score Sums with SAT Scores and High School GPA

Variable	Mean	Minimum	Maximum	Pearson Correlation of Writing Assessment Sum of Scores with SAT Scores
Sum score	16.5709	5	27.5	
SAT Essay	8.60448	6	12	0.08834
SAT Verbal	589.02985	390	800	0.21197*
SAT Writing	583.50746	340	740	0.20483*
High School GPA	3.10082	2.15	4	0.28939***

* significant at $\leq .05$

**significant at $\leq .01$

***significant at $\leq .001$

Second Phase, 2007-2008, of Writing Program Assessment: Methodology

- ❖ After phase I, writing faculty wanted to revise the assessment process. Some wanted more qualitative information about students' learning about the writing process.
- ❖ IR provided feedback both on prompt and rubric. Three student responses were chosen to use for training faculty.
- ❖ Email accounts were set up just for the purpose of assessment where students emailed their responses to the prompt. IR retrieved all prompts and erased any identifiers e.g. names, and course number etc.
- ❖ In early May, 2008, ten writing faculty graded 134 writing prompts for half a day followed by discussions of their grading experiences.

Methodology [Cont'd]

- ❖ In May 2008, 134 papers from four spring 2008 freshmen writing courses were evaluated by the English writing faculty. The break down by course is as follows:
 - ❖ ENG 104[Academic Writing I]: 12 [22% of total course enrollment of 59]
 - ❖ ENG 105[Academic Writing II]: 117 [42.1% of total course enrollment of 278]
 - ❖ ENG 106[Academic Writing III]: 5 [20% of total course enrollment of 25]
 - ❖ **Total of all classes: 134 [37% of 366]**

May 2008 : Instructions for Writing Prompt

Dear Student,

This semester, as part of the college's Writing Program assessment, we are requesting that all members of English 103, 104, 105, and 106 courses write a short response to a common prompt.

Please respond to the prompt in detail and be as articulate as possible. The greater the level of specificity in your response, the more useful your personal point of view will be to us. Be mindful of your grammar and mechanics, and, most importantly, be honest.

Your responses will help the Writing Program to facilitate the growth and development of student writing. This is not a graded assignment. Your course instructor will not see your response, and it will not have any impact on your grade for the course.

Please email your completed response to ENG104@goucher.edu by March 10th. Please send it as an attachment in .doc format.

May 2008 : Writing Prompt

Write a 1 to 2 page description of your writing process.
Be sure to include the following points:

1. What steps do you take, from start to finish, to complete a college-level writing assignment?
2. Are you usually satisfied with the results of this process? Why or why not?

Writing Assessment Rubric for Prompt

Part I: Circle the most applicable option

Writing Process

- The student describes a detailed writing process (3).
- The student describes a writing process that is vague or generalized(2).
- The student does not describe a writing process or admits to having little or no process(1).

Presentation

- The student's writing is clear and precise(3).
- The student's writing is clear but uneven(2).
- The student's writing is riddled with errors(1).

Investment

- The student demonstrates a strong investment in her written work(3).
- The student demonstrates some investment in her written work(2).
- The student does not demonstrate investment in her written work(1).

Writing Assessment Rubric for Prompt [cont'd]

Part II: Check all that apply

___The student admits to wanting to **spend more time** on her writing

___The student **utilizes The Writing Center** and/or peer review.

___The student discusses **grammar and mechanics**.

___The student describes a process that includes **prewriting strategies** (outlining, prewriting, etc.).

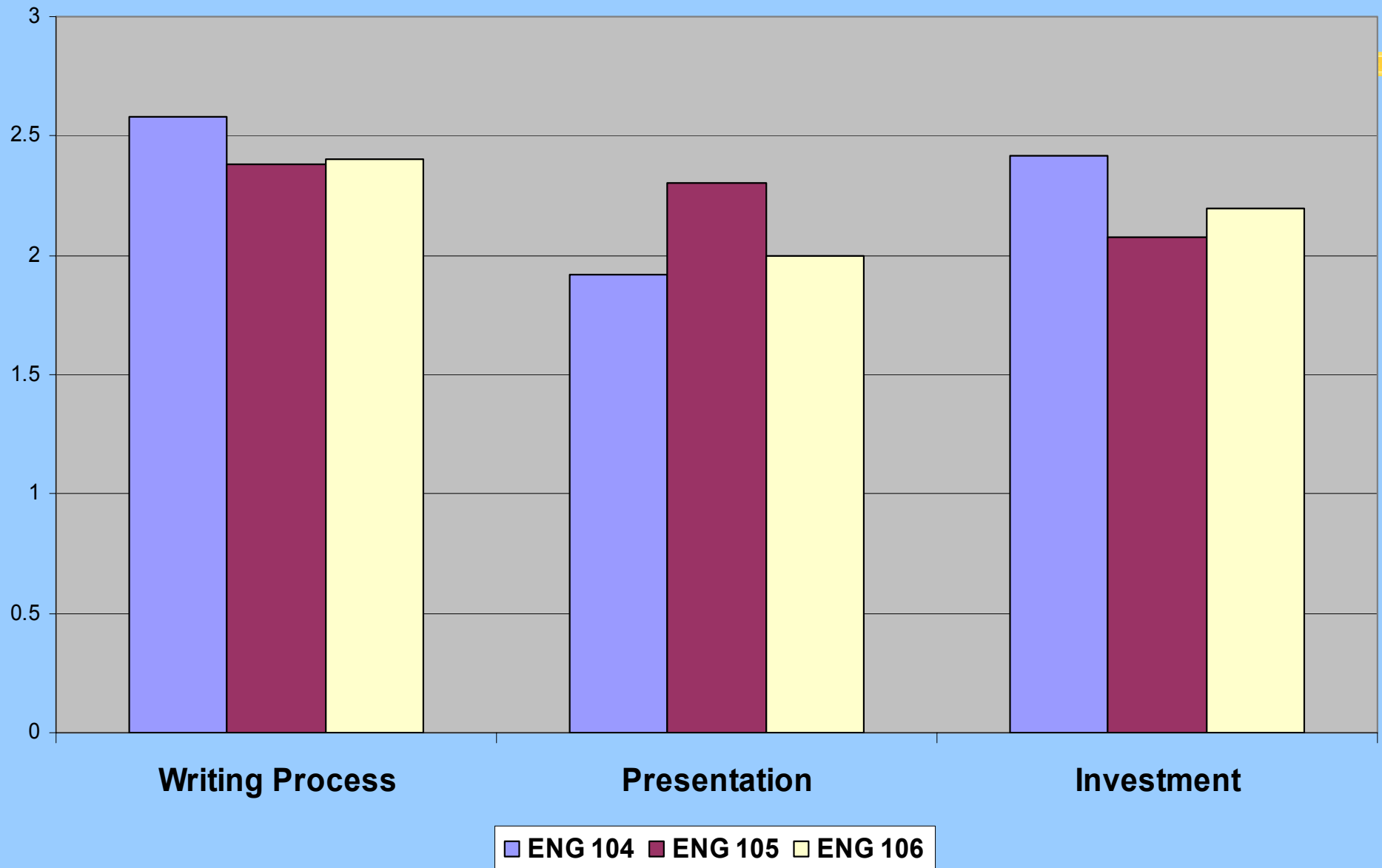
___The student describes a process that includes **multiple drafts**.

___The student describes a process that incorporates **instructor feedback**.

___The student shows **awareness of audience**.

Part III: Additional Comments (use back of page if necessary)

Spring '08 Writing Program Assessment: Learning Outcomes Mean Scores by Course



**Writing Strategies Used by Students
By English Writing Class: Spring, 2008**

Rubric Components	English 104		English 105		English 106	
	n	%	n	%	n	%
wanting to spend more time on her writing.	2	16.67%	21	17.95%	1	20.00%
utilizes Writing Center and/or peer review	6	50.00%	46	39.32%	3	60.00%
discusses grammar and mechanics.	4	33.33%	44	37.61%	2	40.00%
includes prewriting strategies.	7	58.33%	78	66.67%	2	40.00%
includes multiple drafts.	6	50.00%	53	45.30%	1	20.00%
incorporates instructor feedback.	3	25.00%	31	26.50%	1	20.00%
awareness of audience.	3	25.00%	19	16.24%	1	20.00%
Total students who wrote a prompt.	12		117		5	

Challenges

- ❖ Stimulating writing faculty to do systematic assessment without any incentives (monetary, time, or symbolic)
- ❖ Resolving differences and achieving consensus among faculty on topics such as rubrics
- ❖ Selecting suitable papers for training faculty for assessment grading
- ❖ Small number of minority students makes it harder to find statistically significant differences in mean scores
- ❖ Having students comply with the request of writing for assessment purposes—since it was not for credit in phase 2.

Thank You

Please email gwisan@goucher.edu
or pallabi.roy@goucher.edu for more info



QUESTIONS?

SUGGESTIONS?

COMMENTS?