

2008 DELNA Screening Pilot

*Nana Lowell and Debbie E McGhee
February 2009*

SUMMARY

In fall of 2007, Provost Wise established an English Language Proficiency Working Group to examine UW English language proficiency policies relating to entering and current students. Over the course of six months, the Working Group considered a variety of issues and alternatives, and developed a set of recommendations to improve UW practices in assessing English language proficiency and providing academic support for students.

A primary concern in forming the Working Group was the University policy requiring that non-citizen students take a diagnostic test of English language proficiency unless they could submit standardized test scores exceeding specified cut scores. The Working Group determined that a more equitable policy would require English language screening of all students, regardless of citizenship. Students who were unable to submit standardized test scores above the specified cuts would be given the opportunity to pass a screening test. As under the current system, students who did not pass the screening would go on to take a longer diagnostic test, and those who continued to score poorly would be required to take Academic English Program (AEP) courses.

This report describes the methods and outcomes of a pilot of the Diagnostic English Language Needs Assessment (DELNA) Screening carried out during September and August of 2008. The DELNA Screening is a short (17 minutes) online test consisting of Vocabulary and Speed-Reading subtests. It was administered during selected freshman, transfer, and international student orientation sessions. A total of 1158 completed tests were obtained. Student test scores were combined with demographic and academic data from the UW Student Database to address four research questions identified by the Working Group.

1. *How do different student groups (first-time freshmen, transfer students, International students) perform on the DELNA Screening?*

Overall, students performed significantly better on the DELNA Vocabulary subtest ($Mn = 91.8\%$) than on the Speed-Reading subtest ($Mn = 62.3\%$). Due to the restriction in range of Vocabulary scores, the Speed-Reading subtest showed better discrimination among groups.

The three orientation groups did not differ significantly in their performance on the Vocabulary subtest. However, on the Speed-Reading subtest, freshmen ($Mn = 65.3\%$) scored higher than both transfer students ($Mn = 58.1\%$) and international ($Mn = 39.4\%$) students. The mean difference between transfer and international students was also statistically significant.

The majority (78.7%) of students indicated that they grew up speaking English. These examinees had significantly higher Vocabulary and Speed-Reading scores. U.S. citizens achieved higher scores than non-citizens. Analysis of variance of total score showed that:

- a) among citizens, there was a large difference (50.3 vs. 76.2) by native language status; and
- b) among non-citizens, there also was a large difference (49.5 vs. 63.5) by native language status; and
- c) among those who grew up speaking English, there was a moderately large difference (76.2 vs. 63.5) between citizens and non-citizens; but
- d) among those who did not grow up speaking English, there was no difference (50.3 vs. 49.5) by citizenship status.

Under-represented minority (URM; Native-, African-, Hispanic/Latino-, Hawaiian/Pacific Islander American) students performed significantly worse than non-URM (i.e., White- and Asian American) students on both subtests. The difference between URM and non-URM students held even after controlling for self-reported native language. There was also an effect for student gender within the non-URM group. Specifically, non-URM men scored higher ($Mn = 75.9$) than non-URM women ($Mn = 73.3$).

We investigated whether students who scored poorly on the DELNA Screening were receiving programmatic services, and found that the majority of DELNA Screening examinees who scored below 50% on Speed-Reading were not enrolled in a special program.

2. *How do results of the DELNA Screening compare to those of other common English language assessments (e.g., SAT CR/Verbal, ACT English, TOEFL, IELTS, AEP Diagnostic)?*

DELNA Screening scores were correlated with other performance measures, most of which are test scores submitted as part of the regular UW admissions process. DELNA Screening score was strongly correlated with ACT and SAT language-related scores, indicating that these tests provide comparable information relative to English proficiency. Although scores on the DELNA Screening were also strongly related to the IELTS, this result is only suggestive due to the low number of IELTS scores available. DELNA Screening scores correlated only weakly with overall high school GPA. Among transfer students, no relationship was found between DELNA Screening scores and either overall 2-year transfer GPA or GPA in 2-year composition classes.

A modest correlation was found between DELNA Screening and AEP Diagnostic scores reflecting the restricted range of English proficiency among students for whom AEP scores were available. To further examine the relationship between these tests and the TOEFL, we identified the AEP placement for each of the 56 examinees who also had taken the AEP Diagnostic. Observed scores on both the DELNA Screening and TOEFL-IBT increased incrementally with AEP placement, consistent with the positive relationship among the three test scores.

In comparing performance on a variety of language-related measures, DELNA Screening was the best discriminator between: a) students attending freshman versus international orientation sessions, b) students attending transfer versus international orientation sessions, and c) U.S. citizens and non-citizens. It was equally as effective as ACT Reading in discriminating between native and non-native speakers of English, and it outperformed the TOEFL IBT in that regard.

3. *What cut scores should be used if regular online screening were implemented?*

Three sets of cut scores are required to determine the number of students who would be tested if all incoming students were screened for English language proficiency, with subsequent testing of students who score below a specified level. These are: a) cut scores for each of the standardized tests to determine whether students would be required to take the DELNA Screening, b) a cut score on the DELNA Screening to determine who should take the AEP Diagnostic, and c) cut scores on the AEP Diagnostic to place students into AEP classes. Standardized test cut scores can be adopted from those currently in

place exempting non-citizen students from AEP Diagnostic testing. Similarly, AEP cut scores exist to place students into classes.

The optimal way to establish a cut score on the DELNA Screening would be to 'link' the DELNA total score with scores from the other standardized measures of English language proficiency. Linking is the process that establishes a statistical relationship between different tests and is the basis for developing concordance tables. At present, there is sufficient data only to establish a tentative link between DELNA total score and SAT CR/Verbal (re-centered) score. Preliminary analyses suggested that the UW's present SAT CR/Verbal criterion of 490 corresponded to a DELNA total score of 55. This value was consistent with the analysis noted above of the relationship between total DELNA Screening score and AEP placement. The mean DELNA Screening score among students exempted from ENGL 103 (i.e., demonstrating proficiency) was 51.8. Furthermore, among the 102 DELNA Screening examinees who achieved a 70 or greater on the TOEFL-IBT the mean DELNA Screening total score was 54.0. Taken together, these results suggest a total DELNA Screening score of 55 to demonstrate English language proficiency.

The numbers of students who would be required to take the DELNA Screening and subsequent AEP Diagnostic were estimated by examining test scores presented by the 6,804 undergraduate students who entered the University in Autumn 2007. If students were required to take the DELNA Screening who either 1) did not present standardized test scores or 2) presented scores below those currently in place to exempt non-citizen students from AEP Diagnostic testing, a total of 1,714 (25.2%) students would have been tested (669 freshmen; 451 transfer students without an AA degree; and 583 transfer students with an AA degree). If the DELNA Screening cut-off were set to 55, 423 (24.7%) students would subsequently take the AEP Diagnostic (165 freshmen, 111 transfer students without an AA degree, and 144 transfer students with an AA degree).

4. How do results of the screening relate to subsequent course grades?

Autumn 2008 course credits and grades were extracted from the UW Student Database for pilot participants. The average number of credits attempted was 14.8 and the 1121 students who received at least one numeric grade earned an overall GPA of 3.21. The average GPA in writing (W) and non-writing courses was 3.32 (n = 273) and 3.19 (n = 1110), respectively. Compared to transfer students with AA degrees, freshmen and transfer students without AA degrees earned significantly higher GPAs overall, as well as significantly higher GPAs in writing and non-writing courses.

The DELNA Screening total score correlated modestly with all three grade variables for freshman entrants. Similar coefficients were obtained for transfer students without AA degrees, but for this group the relationship of DELNA total to Writing GPA was not significant due to the small number of students. The DELNA total was not related to the three grade variables for transfer students with AA degrees. Subsequent regression analyses indicated that, given a DELNA total of 55, we would predict that:

- the average first-quarter overall GPA for freshmen and transfer students without AA degrees will be $3.13 \pm .046$;
- the average first-quarter overall GPA for transfer students with AA degrees will be $2.94 \pm .093$;
- the average first-quarter writing GPA for freshmen will be $3.30 \pm .075$;
- the average first-quarter writing GPA for transfer students (with or without AA degrees), will be $3.06 \pm .156$;
- the average first-quarter non-writing GPA for freshmen and transfer students without AA degrees will be $3.12 \pm .049$;
- the average first-quarter non-writing GPA for transfer students with AA degrees will be $2.91 \pm .098$.

BACKGROUND

In fall of 2007, Provost Wise established an English Language Proficiency Working Group to examine several questions relating to the fairness and efficacy of UW English language proficiency policies relating to entering and current students. The Working Group was asked to consider which students should be assessed for proficiency, how they should be assessed, what types of coursework or other support the university should provide, and who should pay for such coursework or support. Over the course of six months, the Working Group considered a variety of issues and alternatives, and developed a set of recommendations to improve UW practices.¹

Important concerns discussed by the group included the current policy requiring English language proficiency testing of all international students but not of native students. The Working Group proposed the alternative policy that all entering students undergo an initial screening for English proficiency, and that the results of that screening be used to identify students who might benefit from a subsequent more detailed assessment. Entrants could meet the screening requirement by presenting a standardized test score or, for those without scores, taking a UW-approved assessment. The Diagnostic English Language Needs Assessment (DELNA) Screening² was identified as potentially useful for this purpose, and the following report describes a pilot administration of this instrument conducted during September and August of 2008. Particular research questions addressed were:

1. How do different student groups (first-time freshmen, transfer students, International students) perform on the DELNA Screening?
2. How do results of the DELNA Screening compare to those of other common English language assessments (e.g., SAT CR/Verbal, ACT English, TOEFL, IELTS, AEP Diagnostic)?
3. What cut scores should be used if regular online screening were implemented?
4. How do results of the screening relate to subsequent course grades?

METHOD

Instrument

The DELNA Screening was identified by Working Group members as a promising new tool to provide quick screening of English language proficiency. It is a short online assessment recently developed at the University of Auckland for use with student populations similar to that of UW, and comprises two sections. Part A (7 minutes) consists of 27 multiple-choice vocabulary items requiring students to identify the best synonym for specified words. In Part B (10 minutes), students are asked to speed read two to three paragraphs (73 lines of text) and to identify an extraneous word on each line. At the University of Auckland, students receive one of three recommendations based on their test results: 1) their academic English language proficiency is appropriate for university study; 2) they should take advantage of existing university centers and offices for support and advice; or 3) they should register for an in-depth diagnostic test to recommend appropriate language support.

Although the instrument has been used extensively with students at the University of Auckland, it is only now being offered for use elsewhere. The developmental nature of the test allowed us to make minor

¹ Report of the English Language Proficiency Working Group, July 10, 2008.

² <http://www.delna.auckland.ac.nz/about.php>

modifications prior to the pilot. In particular, we reviewed all items and text to ensure that they were compatible with American English. We also added items at the beginning of the assessment asking students “What language(s) did you grow up speaking? (English vs. Other)”, “Which language is your best for...casual conversation (English vs. Other)...reading (English vs. Other)...writing (English vs. Other)?” Students were asked to specify an alternate language for any question to which they answered “Other.”

Sample

The DELNA Screening was administered to students who attended selected freshman and transfer summer orientation sessions in August and September 2008, and to all students attending the single international student orientation in September. Orientation dates and the numbers of students tested are shown in Table 1. Of the 1,618 students scheduled, 1,202 (74.3%) reported for testing and 1,158 (71.6%) completed the assessment.

Table 1. Number of students tested

Date	Orientation Type								
	Freshman			Transfer			International		
	Roster	Present	Completed	Roster	Present	Completed	Roster	Present	Completed
Aug 11				62	46	45			
Aug 19	269	233	228						
Aug 21				57	51	47			
Aug 22	272	243	241						
Aug 25				56	49	39			
Aug 26	274	196	188						
Aug 28				55					
Aug 29	250	217	211						
Sep 8				54	46	43			
Sep 11				56	43	40			
Sep 15							213	78	76
Total	1065	889	868	340	235	214	213	78	76

Procedures

During freshman orientation, online testing sessions were conducted in non-sequential, 30 minute intervals throughout the second day of the session in the general access computer lab on the first floor of Mary Gates Hall (MGH 131). Transfer orientation sessions included smaller numbers of students, and complete groups were tested in a single sitting in a dedicated computer lab (MGH 241). International students were tested in the middle of their orientation day, in a large single group, in MGH 131.

RESULTS AND DISCUSSION

UW student numbers were collected at the time of the online screening and used to retrieve information from the UW Student Database. By linking student demographic and academic variables to screening test scores, we were able to carry out a series of analyses to address the research questions cited above.

1. How do different student groups perform on the DELNA Screening?

Overall, students performed significantly better on the Vocabulary subtest ($Mn = 91.8\%$) than on the Speed-Reading subtest ($Mn = 62.3\%$), $F(1,1139) = 279$, $p = 3.4E-56$, as shown in Table 2.

The three orientation groups did not differ significantly in their performance on the Vocabulary subtest. However, there were significant differences in performance on the Speed-Reading subtest. Freshmen students ($Mn = 65.3\%$) scored higher than both transfer students ($Mn = 58.1\%$) and international ($Mn = 39.4\%$) students, $F(2,1139) = 5.39$, $p = .005$. The mean difference between transfer and international students was also statistically significant.

Table 2. Mean DELNA Screening performance by orientation type

DELNA subtest	Orientation Type											
	Freshman			Transfer			International			Total		
	Mean	SD	n	Mean	SD	n	Mean	SD	n	Mean	SD	n
Vocabulary												
Number answered	27.0	.4	855	26.9	1.1	213	27.0	.0	75	27.0	.6	1143
Total correct	24.9	2.1	855	24.5	2.6	213	24.1	2.8	75	24.8	2.3	1143
Speed-Reading												
Number answered	53.8	15.2	855	48.0	17.5	213	38.7	14.1	75	51.8	16.1	1143
Total correct	47.7	17.4	855	42.5	19.7	213	29.0	14.3	75	45.5	18.3	1143
Total												
Number answered	80.8	15.2	855	74.9	17.8	213	65.7	14.1	75	75	78.7	16.2
Total correct	72.6	18.6	855	67.1	21.4	213	53.1	15.6	75	75	70.3	19.6

The majority (78.7%) of students indicated that they grew up speaking English and, of these, very few preferred another language for conversation, reading, or writing (Table 3). However, sizable percentages of non-native English speakers reported preferring English for these activities (32.6%, 34.6%, and 40.6%, respectively).

Table 3. Responses to lead-in language questions among native English speakers (percentages)

		Orientation Type			
		Freshman	Transfer	Internat.	Total
Total		85.9	73.4	10.5	78.7
English best for conversation	Yes	92.6	92.8	37.5	91.8
	No	6.0	4.3	3.3	4.6
English best for reading	Yes	92.3	92.3	42.1	91.3
	No	7.7	4.3	0.0	4.2
English best for writing	Yes	91.8	90.1	28.0	89.9
	No	9.7	7.0	2.0	6.4

As shown in Table 4, examinees who reported that they grew up speaking English had significantly higher Vocabulary and Speed-Reading scores, $F(1,1133) = 82.1, p = 5.5E-19$. This difference was greatest on the Speed-Reading subtest.

Table 4. Mean DELNA Screening performance by self-reported "native" language

DELNA subtest	Spoke English Growing Up					
	No			Yes		
	Mean	SD	n	Mean	SD	n
Vocabulary						
Number answered	26.9	1.2	243	27.0	.2	900
Total correct	23.4	2.9	243	25.1	1.9	900
Speed-Reading						
Number answered	36.0	13.4	243	56.0	14.0	900
Total correct	26.2	13.3	243	50.7	15.8	900
Total						
Number answered	62.9	13.6	243	83.0	14.0	900
Total correct	49.7	14.7	243	75.8	16.9	900

As expected, U.S. citizens achieved higher scores than non-citizens (Table 5). Although the mean differences were statistically significant for both subtests, the effect of citizenship was much greater on Speed-Reading scores.

Table 5. Mean DELNA Screening performance by citizenship status

DELNA subtest	Citizenship Status					
	U.S. citizen			Non-citizen		
	Mean	SD	n	Mean	SD	n
Vocabulary						
Number answered	27.0	.2	928	26.9	1.2	215
Total correct	25.1	2.0	928	23.6	2.9	215
Speed-Reading						
Number answered	55.1	14.6	928	37.5	14.2	215
Total correct	49.6	16.5	928	27.7	14.7	215
Total						
Number answered	82.0	14.6	928	64.3	14.5	215
Total correct	74.7	17.6	928	51.3	16.2	215

Average DELNA Screening scores by both citizenship status and native language are presented in Table 6. In general, native-English speaking status made a bigger difference than citizenship status. The results of a 2x2 analysis of variance of total score showed that:

- e) among those who were citizens, there was a large difference (50.3 vs. 76.2) by native language status, $F(1,1139) = 132.9, p = 3.7E-29$; and
- f) among non-citizens, there also was a large difference (49.5 vs. 63.5) by native language status, $F(1,1139) = 17.8, p = 2.6E-5$; and
- g) among those who grew up speaking English, there was a moderately large difference (76.2 vs. 63.5) between citizens and non-citizens, $F(1,1139) = 16.6, p = 5.0E-5$; but
- h) among those who did not grow up speaking English, there was no difference (50.3 vs. 49.5) by citizenship status, $F < 1$.

Table 6. DELNA Screening performance by U.S. citizenship status and English language status

Citizenship status	Spoke English Growing Up								
	No			Yes			Total		
	Mean	SD	n	Mean	SD	n	Mean	SD	n
Vocabulary									
U.S citizen	86.0	10.0	56	93.2	7.1	872	92.8	7.5	928
Non-citizen	87.1	10.9	187	90.6	9.7	28	87.5	10.8	215
Total	86.8	10.7	243	93.1	7.2	900	91.8	8.5	1143
Speed Reading									
U.S citizen	37.0	20.4	56	69.9	21.2	872	68.0	22.6	928
Non-citizen	35.6	17.6	187	53.4	27.7	28	37.9	20.1	215
Total	35.9	18.3	243	69.4	21.6	900	62.3	25.1	1143
Total Score									
U.S citizen	50.3	16.1	56	76.2	16.5	872	74.7	17.6	928
Non-citizen	49.5	14.3	187	63.5	22.2	28	51.3	16.2	215
Total	49.7	14.7	243	75.8	16.9	900	70.3	19.6	1143

There were no statistically significant main effects for student gender on either subtest (Table 7).

Table 7. DELNA Screening performance by student gender

DELNA subtest	Student Gender					
	Women			Men		
	Mean	SD	n	Mean	SD	n
Vocabulary						
Number answered	26.9	.8	628	27.0	.1	515
Total correct	24.5	2.4	628	25.1	2.1	515
Percent correct	90.9	8.9	628	92.9	7.8	515
Speed-Reading						
Number answered	51.8	16.0	628	51.7	16.2	515
Total correct	45.0	18.2	628	46.1	18.5	515
Percent correct	61.7	24.9	628	63.1	25.3	515
Total						
Number answered	78.7	16.1	628	78.7	16.3	515
Total correct	69.6	19.5	628	71.1	19.7	515

As a group, under-represented minority (URM; Native-, African-, Hispanic/Latino-, Hawaiian/Pacific Islander American) students performed significantly worse than non-URM (i.e., White- and Asian American) students on both subtests (Table 8). There was also an effect for student gender within the non-URM group. Specifically, non-URM men scored higher ($Mn = 75.9$) than non-URM women ($Mn = 73.3$), $F(1,1139) = 6.66$, $p = .01$ (Figure 1).

Table 8. Vocabulary and Speed-Reading performance by Under-Represented Minority status

DELNA subtest	Student Group								
	Non-URM			URM			Other/Unknown		
	Mean	SD	n	Mean	SD	n	Mean	SD	n
Vocabulary									
Number answered	27.0	.7	800	27.0	.2	128	27.0	.0	215
Total correct	25.1	2.1	800	23.9	2.8	128	24.3	2.3	215
Speed-Reading									
Number answered	54.9	15.1	800	47.3	16.3	128	42.6	15.5	215
Total correct	49.5	17.0	800	39.5	17.9	128	34.2	17.7	215

Note. Other/Unknown includes 150 International students and 55 American students of unknown ethnicity.

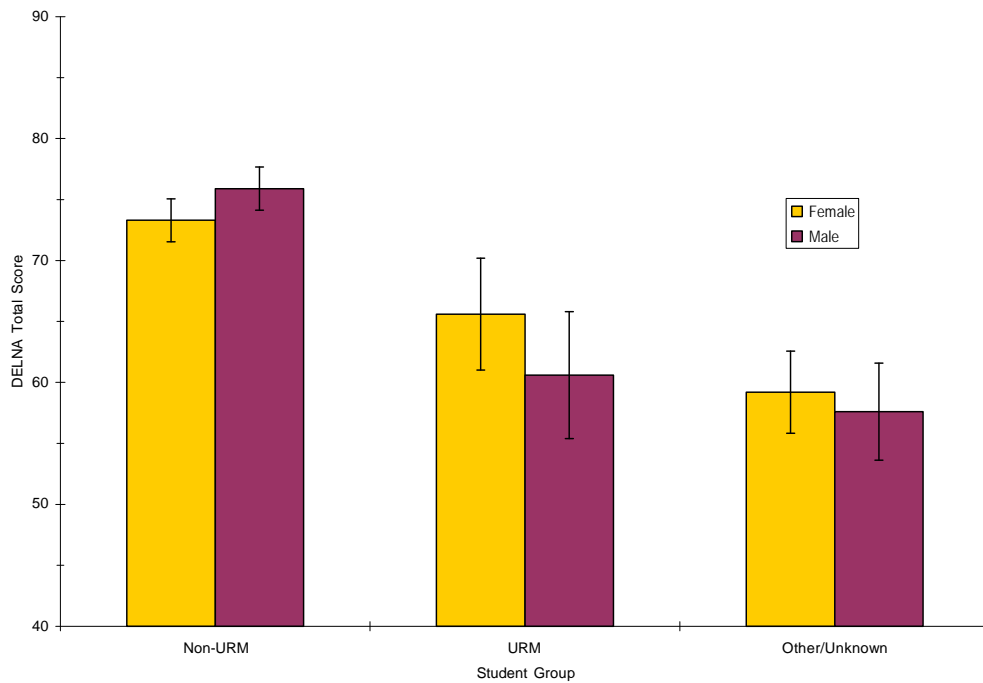


Figure 1. Mean DELNA Screening total scores by student gender and URM status

The difference between URM and non-URM students on DELNA Speed-Reading was observed even after controlling for self-reported native language (Figure 2). For example, among native English speakers, the DELNA Speed-Reading average was 57.9% for URM examinees and 71.1% for non-URM examinees, $F(1,924) = 36.1, p = 2.7E-09, d = -.62$.

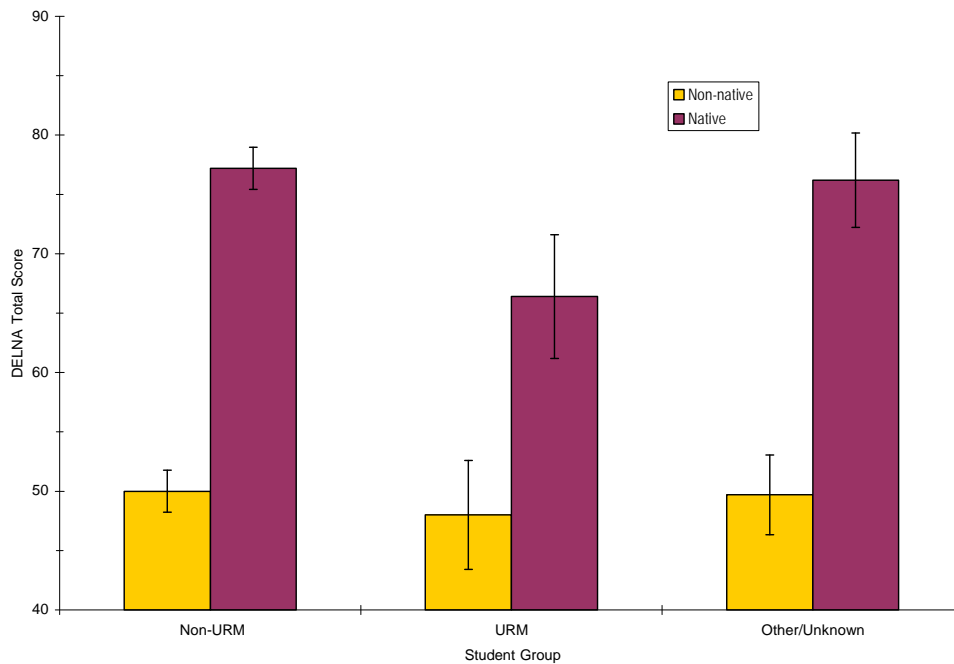


Figure 2. Mean DELNA Screening total scores by English native (L1) and URM status

We investigated whether students who scored poorly on the DELNA Screening were receiving programmatic services. As shown in Table 10, the majority of DELNA Screening examinees who scored below 50% on Speed-Reading were not enrolled in a special program.

Table 9. Special Program involvement among students who scored below 50% on Speed-Reading

Special Program	Orientation Type			Citizenship		English Growing Up	
	Freshman	Transfer	International	U.S. Citizen	Non-citizen	No	Yes
None	164 (68.6)	72 (86.7)	54 (98.2)	169 (67.6)	121 (99.2)	119 (66.9)	171 (85.9)
1: EOP 1	30 (12.6)	9 (10.8)		38 (15.2)	1 (0.8)	32 (18)	7 (3.5)
13: EOP 3	32 (13.4)			32 (12.8)		16 (9.0)	16 (8.0)
30: Athletics	5 (2.1)			5 (2.0)		5 (2.8)	
34: SP Athletics	4 (1.7)			4 (1.6)		4 (2.2)	
60: Staff/Faculty Exempt		1 (1.2)		1 (0.4)			1 (0.5)
93: ADM Committee admit	1 (0.4)			1 (0.4)			1 (0.5)
UWSDB Missing	3 (1.3)	1 (1.2)	1 (1.8)			2 (1.1)	3 (1.5)
Total	239 (100%)	83 (100%)	55 (100%)	250 (100%)	122 (100%)	178 (100%)	199 (100%)

Because all of the analyses thus far had indicated that the Speed-Reading subtest was a much stronger discriminator than the Vocabulary subtest, we performed three multinomial logistic regression analyses in order to test directly whether the Vocabulary score added anything significant over and above the Speed-Reading score. The three dependent variables were orientation type, citizenship status, and self-reported native language. For each analysis, Vocabulary and Speed-Reading scores were entered stepwise. In all three cases, Speed-Reading was a statistically significant predictor of group membership. Vocabulary added only to the prediction of U.S. citizenship status. That is, Vocabulary did not significantly improve the prediction of orientation type, nor did it improve the prediction of self-reported native language beyond what was already predicted by Speed-Reading.

2. How do results of the DELNA Screening compare to those of other common English language assessments?

Table 10 shows the relationships between DELNA Screening scores and other performance measures, most of which are test scores submitted as part of the regular UW admissions process. DELNA Screening scores correlated only weakly with overall high school GPA. Both DELNA Screening subtests were strongly correlated with ACT and SAT language-related scores, indicating that these test scores provide comparable information relative to English proficiency. Although scores on the DELNA Screening were also strongly related to the IELTS, this result is only suggestive due to the low number of IELTS scores available. The modest correlation between DELNA Screening and AEP Diagnostic scores reflects the restricted range of English proficiency among students for whom AEP scores were available.

Table 10. Correlations among DELNA Screening scores and other performance variables

Performance Measure	DELNA Subtest			<i>n</i>
	Vocabulary	Speed-Reading	Total Score	
High School GPA	.13**	.17**	.18**	951
AP Eng Lang Comp	.28**	.47**	.48**	183
ACT English	.57**	.72**	.75**	238
ACT Reading	.53**	.73**	.76**	238
SAT CR/Verbal	.63**	.73**	.76**	908
SAT Writing	.55**	.64**	.66**	883
TOEFL-IBT	.49**	.56**	.60**	135
IELTS	.79*	.73*	.80*	8
AEP Reading/Writing: Comprehension	.32*	.27*	.33*	56
AEP Reading/Writing: Grammar	.36**	.46**	.49**	56
AEP Listening	.34*	.12	.22	55

Note. * Indicates *r* was significantly different from zero with $p < .05$.

** Indicates *r* was significantly different from zero with $p < .01$.

Commonly accepted small, medium, and large *r*-values are .1, .3, and .5, respectively.

To further examine the relationship between the AEP Diagnostic and the DELNA Screening and TOEFL, we identified the AEP placement for each of the 56 examinees who also had taken the AEP Diagnostic. Average scores by placement category are shown in Table 11. Observed scores on both the DELNA Screening and TOEFL-IBT increased incrementally with AEP placement, consistent with the positive relationship among the three test scores. The average total DELNA Screening score among those who were exempt from ENGL 103 was 51.8, suggesting that a DELNA Screening cut score to exempt students from further testing should be at least 50.

Table 11. Mean DELNA Screening and TOEFL scores by AEP Diagnostic Exam placement

Placement	DELNA Screening Performance											
	Vocabulary Total				Speed-Reading Total				TOEFL-IBT Score			
	Min	Mean	SD	<i>n</i>	Min	Mean	SD	<i>n</i>	Min	Mean	SD	<i>n</i>
AEP Writing												
ENGL 100	7	16.4	6.0	5	2	9.2	8.7	5	64.0	64.0	--	1
ENGL 101	13	20.9	3.7	19	7	17.3	6.6	19	35.0	61.4	10.5	9
ENGL 102	16	22.3	2.1	21	10	19.9	6.7	21	49.0	61.3	5.9	10
ENGL 103	19	22.3	2.7	6	14	21.0	5.8	6	58.0	65.5	8.1	4
Exempt from 103	23	24.8	1.5	4	15	27.0	12.8	4	67.0	67.0	--	1
AEP Listening												
ENGL 104	16	19.4	3.3	7	12	17.6	4.4	7	58.0	62.0	5.7	2
Exempt from 104	7	21.7	3.7	47	2	18.7	8.5	47	35.0	62.2	8.2	22

Next we examined whether Transfer GPA or Transfer Composition GPA could be used as a screening mechanism for those transfer students who are also not native English speakers. The Office of Admissions provided data for the 58 DELNA Screening examinees who previously had attended a two-year community or technical college and who indicated that English was not their native language. The average Transfer Overall GPA for this group was 3.56, and the average Transfer Composition GPA was 3.44. Neither overall GPA nor composition GPA correlated with DELNA Screening scores (all $|r| < .08$).

Table 12 provides a summary of standardized group differences in performance on a variety of language-related performance measures. DELNA Speed-Reading was the best discriminator between: a) students attending freshmen versus international orientation sessions, b) students attending transfer versus international orientation sessions, and c) U.S. citizens and non-citizens. It was equally as effective as ACT Reading in discriminating between native and non-native speakers of English; furthermore, it outperformed the TOEFL IBT in that regard. Transfer orientation examinees outperformed international orientation examinees on DELNA Speed-Reading ($d = .74$), but the reverse was true for the TOEFL IBT ($d = -.74$).

Table 12. Standardized effect sizes (d) for group differences on DELNA Screening and other performance measures

Measure	Comparison					
	Freshman / Transfer	Freshman / Int'l	Transfer / Int'l	URM / non-URM	U.S. Citizen / non-citizen	L1 Yes / L1 No
DELNA Vocabulary	.17	.37	.16	-.52	.64	.78
DELNA Speed-Reading	.29	1.11	.74	-.57	1.36	1.61
DELNA Total	.29	1.06	.70	-.61	1.34	1.59
ACT English	.30	--	--	-.74	1.05	1.46
ACT Reading	.48	--	--	-.53	1.27	1.60
SAT CR/Verbal	.00	.60	.57	-.78	.81	1.04
SAT Writing	.02	.31	.30	-.78	.47	.72
TOEFL IBT	.79	.07	-.74	--	--	.76
AEP Reading/Writing: Comprehension	-.18	.22	.33	--	-.30	--
AEP Reading/Writing: Grammar	-.30	.35	.55	--	.09	--
AEP Listening	.33	.66	.22	--	.11	--

Note. Effect sizes were computed only when both cell sizes exceeded $n = 10$. A negative value indicates that the mean of the second group was greater than the mean of first group. Commonly accepted small, medium, and large d -values are .3, .6, and .8, respectively.

3. What cut scores should be used if regular online screening were implemented?

Ultimately, one would like to be able to link DELNA total score with scores from other standardized measures of English language proficiency. Linking is the process by which one establishes a statistical relationship between different tests (SAT and ACT concordance being a prime example). At present, however, our ability to establish a link between the DELNA Screening and other tests is hindered due to the small sample size. At best, given the present data, we can establish only a tentative link between DELNA total score and SAT CR/Verbal (re-centered) score. As shown in Table 13, preliminary analyses suggest that the UW's present SAT CR/Verbal criterion of 490 corresponds to a DELNA total score of 55.

Table 13. Preliminary concordance between DELNA Screening total and SAT CR/Verbal (recentered) score

SAT CR/Verbal (recentered)	DELNA Total	
	Observed Mean ($n = 886$)	Mean-Sigma Concordance
450	51.4	47.2
460	53.4	49.2
470	54.1	51.2
480	53.0	53.2
490	60.0	55.2
500	62.6	57.2
510	62.8	59.2
520	64.5	61.2

Furthermore, among the 102 DELNA Screening examinees who achieved a 70 or greater on the TOEFL-IBT the mean DELNA Screening total score was 54.0. Taking measurement error into account, this result suggests that most English-proficient International students should be able to achieve a DELNA Screening score of 51.2 - 56.8.

Figure 3 shows the proposed decision flowchart for English language proficiency testing. As shown in the figure, all entering students would be screened for English language proficiency either by means of standardized test scores submitted as part of the regular UW admissions process or by the DELNA Screening. Three different sets of cut scores would be identified: a) cut-offs for each of the standardized tests to determine whether students would be required to take the DELNA Screening, b) a cut-off on the DELNA Screening to determine who should take the AEP Diagnostic, and c) cut-offs on the AEP Diagnostic to place students into AEP classes. Transfer students presenting an AA degree and native speakers would be exempted from AEP classes but required to participate in mandatory academic advising.

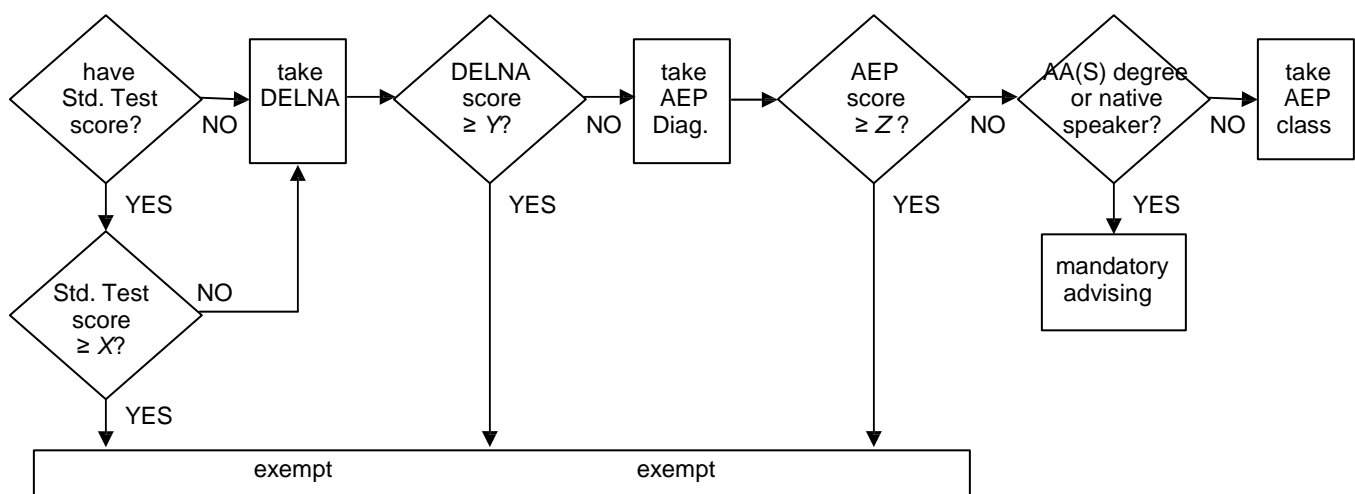


Figure 3. Decision flowchart for English language proficiency screening

Table 14 shows the number of AU 2007 undergraduate entrants who would have been required to take the DELNA Screening because they either 1) did not present standardized test scores, or 2) presented scores below those currently in place to exempt non-citizen students from AEP Diagnostic testing. As shown, 1,714 (25.2%) of the 6,804 incoming students would have been required to take the DELNA Screening (669 freshmen; 451 transfer students without an AA degree; and 583 transfer students with an AA degree). (Note that freshman/transfer classification was missing for eleven students.) Subsequent analyses found similar percentages for undergraduate students who entered during the full 2008 calendar year, as shown in Table 20 in the Addendum.

Table 15 estimates the number of AU 2007 entrants who would have been required to take the AEP Diagnostic test under three different DELNA Screening cut-offs. The table applies DELNA Screening performance rates from the AU 2008 test group to the AU 2007 cohort as a whole. Administering 1,714 DELNA Screening tests and assuming cut-offs of 50, 55, and 60 results in a total number of 316 (18.4%), 423 (24.7%) and 535 (31.2%) students tested, respectively. With the DELNA Screening cut-off set to 55, 165 freshmen, 111 transfer students without an AA degree, and 144 transfer students with an AA degree would be required to take the AEP Diagnostic. Figure 4 is a graphical depiction of the numbers for select student groups.

Table 14. Actual Number of Autumn 2007 undergraduate entrants who satisfied one or more English Language proficiency criteria.

Student Group	AU07 Cohort		No Test Scores		Below Minimum		Above Minimum	
	N	Col. %	N	Row %	N	Row %	N	Row %
Total	6804	100.0	861	12.6	853	12.5	5090	74.8
Gender								
F	3582	52.6	440	12.3	483	13.5	2659	74.2
M	3222	47.4	421	13.1	370	11.5	2431	75.5
URM status								
non-URM	5350	78.6	574	10.7	541	10.1	4235	79.2
URM	755	11.1	103	13.6	192	25.4	460	60.9
other/unknown	699	10.3	184	26.3	120	17.2	395	56.5
U.S. citizenship status								
citizen	5895	86.6	667	11.3	618	10.5	4610	78.2
non-citizen	909	13.4	194	21.3	235	25.8	480	52.8
Freshman	5281		16	.3	653	12.4	4612	87.3
citizen	4785	90.6	8	.2	518	10.8	4259	89.0
non-citizen	496	9.4	8	1.6	135	27.2	353	71.2
Transfer (no degree)	706		345	48.9	106	15.0	255	36.1
citizen	374	53.0	185	49.5	31	8.3	158	42.2
non-citizen	332	47.0	160	48.2	75	22.6	97	29.2
Transfer (AA degree)	806		492	61.0	93	11.5	221	27.4
citizen	727	90.2	466	64.1	68	9.4	193	26.5
non-citizen	79	9.8	26	32.9	25	31.6	28	35.4

Note. UW Registrar's English Language Requirements were: SAT CR/Verbal = 490, ACT = 20, TOEFL-P = 580, TOEFL-C = 237, TOEFL-I = 70, or UW MLT = 90.

TABLE 15. ESTIMATED number of AU 2007 entrants who would have taken AEP Diagnostic under three DELNA total score cut-off scenarios

Student Group	No Score or Not Proficient N	DELNA Total ≥ 50		DELNA Total ≥ 55		DELNA Total ≥ 60	
		DELNA < 50	DELNA ≥ 50	DELNA < 55	DELNA ≥ 55	DELNA < 60	DELNA ≥ 60
		Take AEP \hat{N}	Exempt \hat{N}	Take AEP \hat{N}	Exempt \hat{N}	Take AEP \hat{N}	Exempt \hat{N}
Total	1714	316	1398	423	1291	535	1179
Gender							
F	923	170	753	228	695	288	635
M	791	146	645	195	596	247	544
URM status							
non-URM	1115	206	909	275	840	348	767
URM	295	54	241	73	222	92	203
Other/Unknown	304	56	248	75	229	95	209
U.S. citizenship							
citizen	1285	237	1048	317	968	401	884
non-citizen	429	79	350	106	323	134	295
Freshman	669	123	546	165	504	209	460
citizen	526	97	429	130	396	164	362
non-citizen	143	26	117	35	108	45	98
Transfer (no degree)	451	83	368	111	340	141	310
citizen	216	40	176	53	163	67	149
non-citizen	235	43	192	58	177	73	162
Transfer (AA degree)	585	108	477	144	441	183	402
citizen	534	99	435	132	402	167	367
non-citizen	51	9	42	13	38	16	35

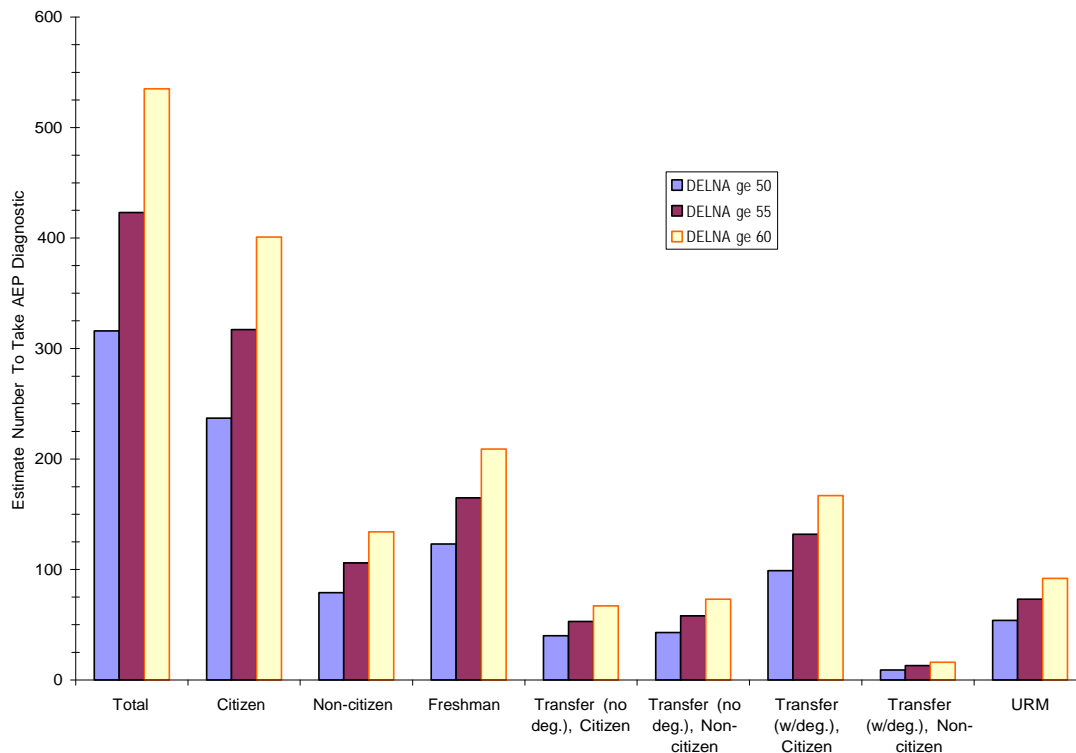


Figure 4. Estimated number of students tested under various scenarios

4. How do results of the DELNA Screening relate to subsequent course grades?

To determine whether student performance on the DELNA Screening was related to subsequent course grades, we extracted autumn 2008 course credits and grades from the UW Student Database for the 1143 pilot participants in the analysis sample. There were no transcript data for fourteen students. The average number of credits attempted was 14.8 (SD = 2.8). For each of the 1121 students who received at least one numeric grade, we computed three grade point averages: overall (GPA), in writing (W) courses (GPA-W), and in non-writing courses (GPA-N). Descriptive statistics are shown in Table 17.

Compared to transfer students with AA degrees, freshmen and transfer students without AA degrees earned significantly higher overall GPAs ($F(2,1118) = 9.25, p = 1.0E-04$), significantly higher writing GPAs ($F(2,270) = 3.28, p = .04$), and significantly higher non-writing GPAs ($F(2,1107) = 9.81, p = 6.0E-05$).

Table 17. Autumn quarter credits and GPAs.

All	Overall			Writing			Non-writing		
	Mean	SD	n	Mean	SD	n	Mean	SD	n
Graded credits	13.1	3.2	1121	5.8	2.1	273	11.8	3.9	1110
GPA	3.21	.62		3.32	.64		3.19	.66	
Freshman									
Graded credits	13.2	3.1	909	5.8	2.1	223	11.9	3.7	899
GPA	3.24	.57		3.37	.63		3.23	.60	
Transfer, no AA									
Graded credits	13.4	3.2	92	6.35	2.4	20	12.1	3.87	91
GPA	3.17	.69		3.18	.86		3.16	.70	
Transfer, AA degree									
Graded credits	12.4	3.7	120	5.67	1.7	30	11.0	4.0	120
GPA	2.98	.83		3.08	.48		2.95	.92	

The DELNA total score correlated modestly with all three grade variables for the sample as a whole, and for freshman entrants (Table 18).³ The coefficients likely underestimate the true strength of the relationships because of range restriction in both DELNA Screening total and (especially) grades. Similar coefficients were obtained for transfer students without AA degrees, but for this group the relationship of DELNA total to Writing GPA was not significant due to the small number of students.

Relationships between DELNA total score and the three grade variables were all non-significant for transfer students with AA degrees.

Table 18. Correlations (Spearman's *rho*) between DELNA and Autumn GPA.

Student Group	All		Freshman		Transfer, no AA		Transfer, AA degree	
	<i>rho</i>	<i>n</i>	<i>rho</i>	<i>n</i>	<i>rho</i>	<i>n</i>	<i>rho</i>	<i>n</i>
All								
Overall GPA	.22**	1121	.22**	909	.23*	92	.14	120
GPA Writing	.28**	273	.28**	223	.26	20	.01	30
GPA Non-writing	.21**	1110	.21**	899	.22*	91	.11	120

Note. * Indicates *r* was significantly different from zero with $p < .05$.

** Indicates *r* was significantly different from zero with $p < .01$.

In order to specify the relationship between DELNA total score and first-quarter grades, we conducted a series of linear regression analyses. Because of the significant differences in the relationships among GPAs and DELNA total score between transfer students with AA degrees and other students, we included student group as a predictor variable. The results are shown in Table 19.

Table 19. Regression coefficients for predicting GPA from DELNA Screening total score.

Student Group	β -Constant	β -Slope
Freshman		
Overall GPA	2.772	.0065
GPA Writing	3.057	.0045
GPA Non-writing	2.767	.0061
Transfer, no AA		
Overall GPA	2.772	.0065
GPA Writing	3.057	0
GPA Non-writing	2.767	.0061
Transfer, AA degree		
Overall GPA	2.772	.0031
GPA Writing	3.057	0
GPA Non-writing	2.767	.0026

For the purpose of predicting future student outcomes, the results of the regression analyses can be interpreted as follows. The general form of any regression equation is:

$$GPA = \beta_{const.} + \beta_{slope} \times DELNAScore$$

³ We computed Spearman's *rho* rather than the more common Pearson's correlation coefficient due to the marked skew in grade data.

Thus, given a DELNA Screening total score of 55, we would predict that:

- for all freshmen and transfer students without AA degrees, their average first-quarter overall GPA will be $3.13 \pm .046$;
- for all transfer students with AA degrees, their average first-quarter overall GPA will be $2.94 \pm .093$;
- for all freshmen, their average first-quarter Writing GPA will be $3.30 \pm .075$;
- for all transfer students (with or without AA degrees), their average first-quarter Writing GPA will be $3.06 \pm .156$;
- for all freshmen and transfer students without AA degrees, their average first-quarter non-writing GPA will be $3.12 \pm .049$;
- for all transfer students with AA degrees, their average first-quarter non-writing GPA will be $2.91 \pm .098$.

Figures 5 through 7 provide graphical depictions of the regression results.

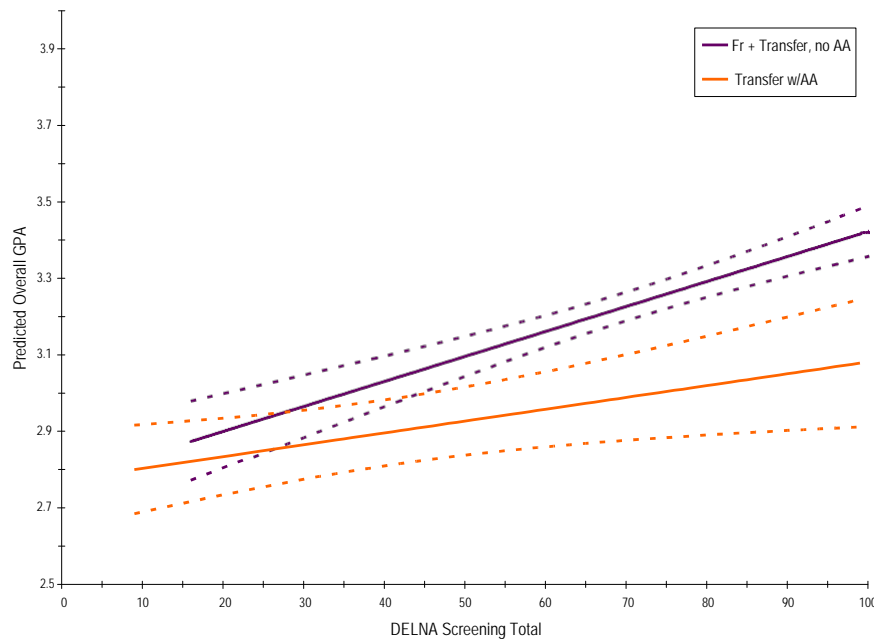


Figure 5. Predicted Overall GPA as a function of DELNA Screening Score (with 95% confidence intervals)

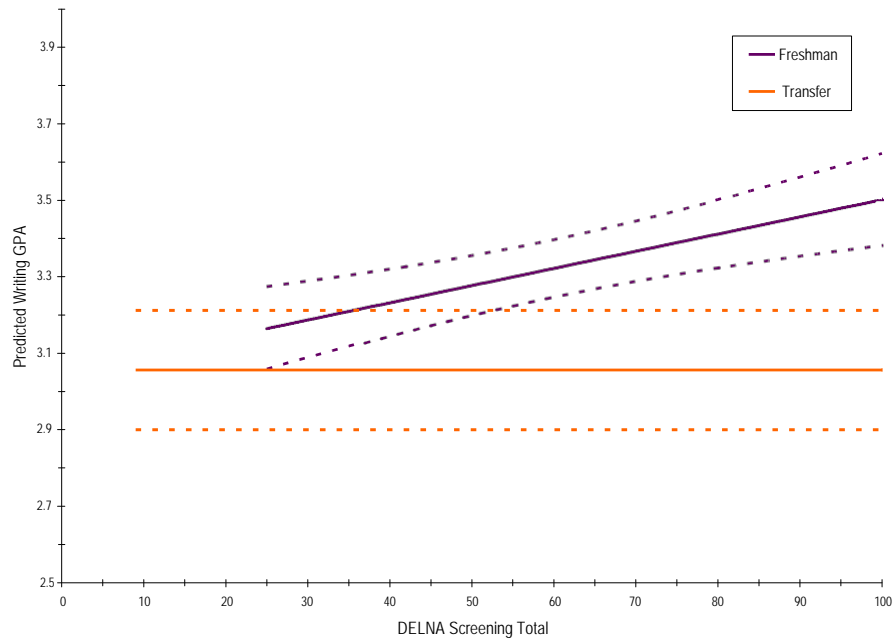


Figure 6. Predicted Writing GPA as a function of DELNA Screening Score (with 95% confidence intervals)

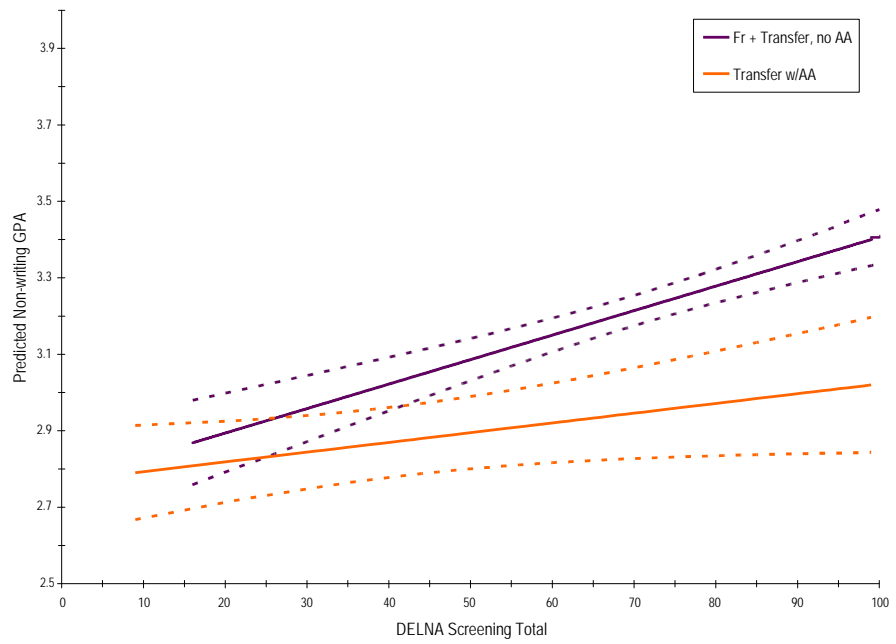


Figure 7. Predicted Non-writing GPA as a function of DELNA Screening Score (with 95% confidence intervals)

ADDENDUM

Table 20. ACTUAL number of Autumn-Winter 2008 undergraduate entrants who satisfied one or more English Language proficiency criteria.

Student Group	Total	No Test Scores		Below Minimum		Above Minimum		
	N	N	Row %	N	Row %	N	Row %	
Gender								
F	4411	827	18.7	497	11.3	3087	70.0	
M	3934	724	18.4	420	10.7	2790	70.9	
URM status								
non-URM	6304	1077	17.1	552	8.8	4675	74.2	
URM	955	215	22.5	218	22.8	522	54.7	
other/unknown	1086	259	23.8	147	13.5	680	62.6	
U.S. citizenship status								
citizen	7101	1287	18.1	627	8.8	5187	73.0	
non-citizen	1244	264	21.2	290	23.3	690	55.5	
WI08								
freshman	22	0	.0	4	18.2	18	81.8	
transfer	291	129	44.3	21	7.2	141	48.5	
transfer AA(S)	330	200	60.6	34	10.3	96	29.1	
total	643	329	51.2	59	9.2	255	39.7	
SP08								
freshman	17	5	29.4	2	11.8	10	58.8	
transfer	163	82	50.3	17	10.4	64	39.3	
transfer AA(S)	141	91	64.5	11	7.8	39	27.7	
total	321	178	55.5	30	9.3	113	35.2	
SU08								
freshman	38	1	2.6	13	34.2	24	63.2	
transfer	145	96	66.2	5	3.4	44	30.3	
transfer AA(S)	96	64	66.7	8	8.3	24	25.0	
total	279	161	57.7	26	9.3	92	33.0	
AU08								
freshman	5540	8	.1	650	11.7	4882	88.1	
transfer	724	414	57.2	52	7.2	258	35.6	
transfer AA(S)	838	461	55.0	100	11.9	277	33.1	
total	7102	883	12.4	802	11.3	5417	76.3	
Total								
freshman	5617	14	.2	669	11.9	4934	87.8	
transfer	1323	721	54.5	95	7.2	507	38.3	
transfer AA(S)	1405	816	58.1	153	10.9	436	31.0	
total	8345	1551	18.6	917	11.0	5877	70.4	

Note. UW Registrar's English Language Requirements were: SAT CR/Verbal = 490, ACT = 20, TOEFL-P = 580, TOEFL-C = 237, TOEFL-I = 70, IELTS = 7, or UW MLT = 90.