

# Assessment of Written Communication and Critical Thinking at BYU-Hawaii

Winter 2016-17 Results Summary by the Office of Institutional Research

**Assessment results for Written Communication in winter 2016-17 are similar to those from the Written Communication assessment completed in winter 2015-16.**



"Writing" by jeffrey james pacres is licensed under CC BY-NC-ND 2.0

Artifacts at both the sophomore and senior levels received the largest proportion of scores at the “developed” stage for most criteria while the separation of scores between these levels narrowed in 2016-17.

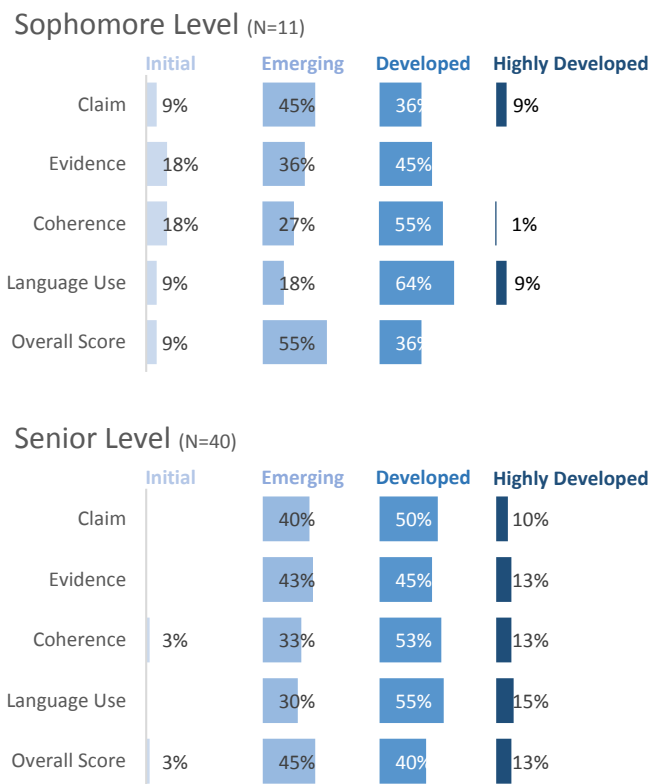
As shown in Chart 1, the largest proportion of artifacts at the sophomore level were at the “developed” stage for all criteria except Claim, while the largest proportion of artifacts at the senior level were at the “developed” stage for all criteria. In addition, the largest proportion of artifacts at both the sophomore and senior levels for Overall Score were at the “emerging” stage.

Artifacts at the senior level had a higher proportion at the “developed” and “highly developed” stages than those at the sophomore level. Conversely, artifacts at the sophomore level had a higher proportion at the “initial” stage than those at the senior level. However, a comparison of means between these groups showed no significant differences for any of the criteria or overall score between sophomores and seniors.

These findings are similar to those from the written communication assessment conducted in winter 2015-16 where the largest proportions of artifacts fell under similar stages for both the sophomore and senior levels. Furthermore, in 2015-16 there was evidence of improvement at the senior level over the sophomore level for the Sources criterion and Overall Score. As stated above, the results from 2016-17 do not show evidence of such improvement.

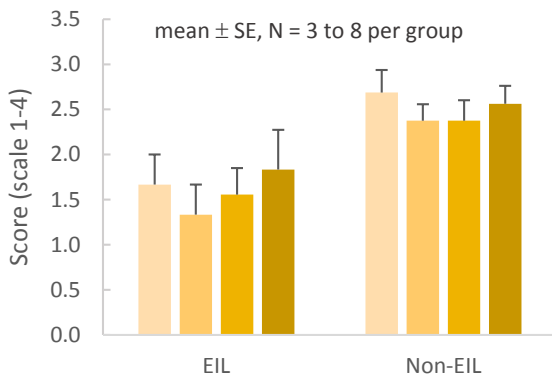
The small number of sophomore level artifacts (N=11) as well as the smaller number of artifacts assessed overall (N=51) could be a contributing reason for these results.

**Chart 1: Sophomores and seniors are at the “developed” stage for most criteria**



**Chart 2: Comparison of Written Communication criteria at sophomore and senior levels**

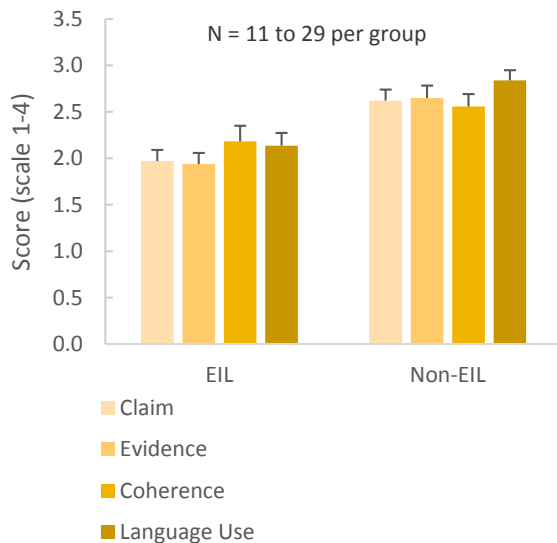
Sophomore Level



There continues to be a consistent difference between the papers written by EIL students to non-EIL students, and this gap widens between the sophomore and senior levels.

There is a significant difference in Overall Scores for the papers written by EIL students versus non-EIL students at both levels, and this difference widens between the sophomore level ( $p < .05$ ) and senior level ( $p < .01$ ). At the sophomore level the major difference is in the Evidence criterion ( $p < .05$ ), while at the senior level non-EIL artifacts scored significantly higher for Claim ( $p < .01$ ), Evidence ( $p < .01$ ), and Language Use ( $p < .001$ ). These differences can be seen in Chart 1.

Senior Level



Although Chart 2 also shows a visible improvement in scores for papers written by EIL students between the sophomore and senior levels, the difference is not statistically significant with this sample. In addition, there is only modest improvement in scores for papers written by non-EIL students between the sophomore and senior levels. These findings are consistent with the results from the 2015-16 assessment of written communication.

There were no statistical differences between scores by gender or major college. And any statistical differences detected between ethnicity and home area were all attributable to EIL status. See Table 3 and Table 4 for more information on the mean scores breakdown for these variables.

Inter-Rater Reliability

Inter-Rater reliability is calculated using a discrepancy index as shown in Table 1. The largest discrepancy between raters was for the Coherence criterion, and general agreement (discrepancy of 0) is in the fair to medium range.

An attempt at calculating Cohen’s Kappa found that there were six different combinations of reader pairs for artifacts ranging in number from 5 to 11

**Table 1: Discrepancy Index for Written Communication**

	DISCREPANCY OF			
	0	1	2	3
Claim	45%	41%	14%	0%
Evidence	49%	41%	10%	0%
Coherence	29%	57%	12%	2%
Language Use	39%	49%	12%	0%

**Figure 1: Criteria for the 2015-16 and 2016-17 versions of the Written Communication rubric**

2015-16	2016-17
<p><b>CONTENT</b> Writing displays effective communication through task completion, attention to context, analysis, and creativity.</p> <p><b>COHERENCE</b> Writing displays appropriate organization related to formatting, paragraphing, purpose, and transitioning.</p> <p><b>LANGUAGE USE</b> Writing follows linguistics conventions such as spelling, punctuation, grammar, and word choice.</p> <p><b>SOURCES &amp; EVIDENCE</b> Writing is appropriately supported by relevant reasons or examples, appropriate citations or source use, and disciplinary conventions.</p>	<p><b>CLAIM</b> - Task completion - Critical, creative, analytic thinking - Awareness of context and multiple perspectives</p> <p><b>EVIDENCE</b> - Logical Analysis - Critical thinking - Critical evaluation of sources and evidence - Relevant, effective support</p> <p><b>COHERENCE</b> Logical, clear organization</p> <p><b>LANGUAGE USE</b> Conventions of spelling, punctuation, grammar, word choice, and tone</p>

papers for each pair. Although overall agreement was fairly high for several reader pairs in the criteria of Evidence and Claim (60-78% agreement) – as also corroborated in Table 2 – the Cohen’s Kappa for some of these pairs was much smaller (after taking into account the amount of agreement that would happen by chance.)

In order to effectively employ other measures of inter-rater reliability (e.g., Cohen’s Kappa, correlation between paired readers), it is recommended that, in the future, at least 25-30 of the artifacts be read by the same two raters.

Essentially, there are not enough artifacts for each reader pair to calculate the inter-rater reliability with a high degree of confidence in this sample. This can be improved upon in the next assessment round for Written Communication.

**Rubric Adjustments**

The Written Communication ILO faculty group adjusted the Written Communication rubric to incorporate elements of Critical Thinking. For this reason the criterion and descriptions have changed from the 2015-16 version and these changes are outlined in Figure 1. Some of the differences between the 2015-16 and 2016-17 assessments could also be attributable to these changes. The rubric version used in the winter 2016-17 assessment was finalized in December 2016 and is attached to this report.

**Observations on Methodology**

A stratified random sample was used to select student artifacts from GE 110, ENGL 201 at the sophomore level, and from ENGL 315, BIOL 494, CHEM 494, and HIST 490 at the senior level. Table 2 on page 5 shows that the demographic proportions of the population are fairly well represented in this sample. Most of the artifacts used are taken from the GE courses (GE 110, ENGL 201, ENGL 315), with

the few remaining coming from senior capstones in programs that do not require ENGL 315.

The faculty group for Information Literacy held a calibration session before the full norming session. During the norming session two separate readers rated each paper and where there was disagreement greater than one whole point on any of the criteria a third reader was employed. The final score is found by taking the average of all readers.

This experience has garnered the following observations that will be helpful in guiding future efforts to assess Information Literacy at BYU-Hawaii.

***Continuous assessment in small batches***

Small sample size is a likely contributor to these results, especially for the sophomore level. However, it is taxing on faculty members to assess large numbers of artifacts. In order to gain the advantage of a larger sample and not overburden faculty members, assessment for Information Literacy could be conducted each year in small batches and then combined for overall analysis. To

do this it is imperative that the same methodology and rubric be used for each assessment session.

***Are the artifacts appropriate?***

Assessment of learning informs instruction, but are these the appropriate artifacts to examine for discovering the learning of information literacy? There is heavy dependence on General Education courses that students may or may not be taking at their sophomore and senior years. Especially at the senior level, it would be well for the institution to consider assessing assignments other than ENGL 315, such as program capstones, that may be better evidence of a student's best effort.

***Consider the instructors***

Is the instructor a contributing factor in these results? It behooves the university to examine how variables such as the instructor's faculty status or length of time at the institution might affect these scores. This information will be useful for faculty coaches in the Center for Learning and Teaching to help begin conversations on English language learners and writing, as well as conversations that can help inform improvement of pedagogy.

### Sample and Representativeness

The sample and population proportions listed in Table 2 show that the sample is fairly representative of the population for most demographic categories. The population is based on Fall 2016 degree-seeking enrollment for all demographic groupings except level. Level (sophomore/senior) is based on the proportion of associates (sophomore level) and bachelors (senior level) degrees that were awarded during the 2015-16 academic year.

**Table 2: Demographic proportions in the sample fairly well represent those of the population**

	SAMPLE N=51	POPULATION N=2601
<b>Gender</b>		
Male	35%	41%
Female	65%	59%
<b>EIL Status</b>		
Enrolled in EIL	28%	33%
Did not enroll in EIL	73%	67%
<b>Ethnicity</b>		
American Indian/Alaska Native	0%	1%
Asian	23%	29%
Black	0%	1%
Hawaiian	13%	4%
Hispanic	5%	6%
Pacific Islander	20%	20%
White	40%	39%
<b>Home Area</b>		
Asia	20%	25%
Pacific	15%	15%
Hawaii	10%	10%
US Mainland	50%	46%
Other International	5%	4%
<b>College</b>		
Arts & Humanities	8%	17%
Business, Computing & Gov't.	33%	36%
Human Development	20%	15%
Math & Sciences	28%	23%
Special Programs	10%	4%
Undecided	3%	5%
<b>Level</b>		<b>N=769</b>
Sophomore (Associates)	22%	23%
Senior (Bachelors)	78%	77%

Mean Scores

Mean scores for each criterion and the overall score are listed by demographic variable grouping for sophomores in Table 3 and seniors in Table 4. The criterion with the highest mean for each is group is highlighted.

**Table 3: Mean scores – Sophomores (N=11)**

Gender	CLAIM	EVIDENCE	COHERENCE	LANGUAGE USE	OVERALL SCORE
Male	2.50	2.00	2.29	2.50	2.32
Female	2.36	2.14	2.07	2.29	2.21
EIL Status	CLAIM	EVIDENCE	COHERENCE	LANGUAGE USE	OVERALL SCORE
Enrolled in EIL	1.67	1.33	1.56	1.83	1.60
Did not enroll in EIL	2.69	2.38	2.38	2.56	2.50
Ethnicity	CLAIM	EVIDENCE	COHERENCE	LANGUAGE USE	OVERALL SCORE
Asian	1.67	1.83	1.67	2.00	1.79
Hawaiian	2.50	2.00	2.50	2.50	2.38
Hispanic	2.00	2.50	2.50	2.50	2.38
Pacific Islander	2.00	1.25	1.33	1.75	1.58
White	3.25	2.63	2.75	2.88	2.88
Home Area	CLAIM	EVIDENCE	COHERENCE	LANGUAGE USE	OVERALL SCORE
Asia	1.67	1.83	1.67	2.00	1.79
Pacific	2.00	1.25	1.33	1.75	1.58
Hawaii	3.00	2.00	3.00	2.50	2.63
US Mainland	2.90	2.60	2.60	2.80	2.73
Other International	-	-	-	-	-
College	CLAIM	EVIDENCE	COHERENCE	LANGUAGE USE	OVERALL SCORE
Arts & Humanities	1.00	1.00	1.00	1.00	1.00
Business, Computing & Gov't.	2.67	2.33	2.50	2.67	2.54
Human Development	2.33	2.00	2.22	2.33	2.22
Math & Sciences	2.17	2.00	1.83	2.17	2.04
Special Programs	4.00	3.00	3.00	3.50	3.38
Undecided	-	-	-	-	-

**Table 4: Mean scores – Seniors (N=40)**

Gender	CLAIM	EVIDENCE	COHERENCE	LANGUAGE USE	OVERALL SCORE
Male	2.37	2.38	2.48	2.71	2.37
Female	2.48	2.49	2.44	2.61	2.48
EIL Status	CLAIM	EVIDENCE	COHERENCE	LANGUAGE USE	OVERALL SCORE
Enrolled in EIL	1.97	1.94	2.18	2.14	2.06
Did not enroll in EIL	2.62	2.65	2.56	2.84	2.67
Ethnicity	CLAIM	EVIDENCE	COHERENCE	LANGUAGE USE	OVERALL SCORE
Asian	2.28	2.24	2.52	2.50	2.38
Hawaiian	2.10	2.50	2.20	3.00	2.45
Hispanic	2.25	2.00	2.08	2.25	2.15
Pacific Islander	2.27	2.15	2.25	2.35	2.26
White	2.75	2.77	2.65	2.81	2.74
Home Area	CLAIM	EVIDENCE	COHERENCE	LANGUAGE USE	OVERALL SCORE
Asia	2.19	2.02	2.33	2.31	2.21
Pacific	2.03	1.94	2.00	2.14	2.03
Hawaii	2.00	2.63	2.25	2.88	2.44
US Mainland	2.83	2.83	2.79	2.95	2.85
Other International	1.75	1.58	1.33	2.00	1.67
College	CLAIM	EVIDENCE	COHERENCE	LANGUAGE USE	OVERALL SCORE
Arts & Humanities	2.61	2.89	2.72	2.83	2.76
Business, Computing & Gov't.	2.17	2.27	2.24	2.50	2.29
Human Development	2.44	2.13	2.19	2.48	2.31
Math & Sciences	2.65	2.73	2.80	2.86	2.76
Special Programs	2.75	2.63	2.50	2.75	2.66
Undecided	2.00	2.50	2.50	2.50	2.38

## Written Communication and Critical Thinking Rubric for BYU Institutional Learning Outcomes, Dec 2016

Attribute	1 (Initial)	2 (Emerging)	3 (Developed)	4 (Highly Developed)
CLAIM task completion critical, creative, analytic thinking awareness of context and multiple perspectives	Lacks a clear purpose and audience.	The position is simplistic, not analytic, or narrow.  Some of the terms of the thesis or position may be unclear.  The essay has a vague sense of audience or context for the argument.	Makes a significant argument.  Terms of the thesis or position are clear.  Has a fairly clear audience or context.  Shows some awareness of other perspectives and limitations of own position.	Clear thesis that responds thoughtfully to the complexity of the question or issue.  The student may carefully position their work in relation to scholarly or professional inquiry.  Responds to potential weaknesses in, or objections to, own viewpoint or offers thoughtful consideration of other perspectives.
EVIDENCE logical analysis critical thinking critical evaluation of sources and evidence relevant, effective support	Support is minimal, irrelevant, ineffective, confusing, inaccurate or biased.  Ineffective, inadequate or confusing citation	Support for some claims or points is occasionally irrelevant or illogical or biased.  Slight, minimal, or inaccurate comment on data, quotation, and information from sources.  Citation shows several errors that interfere with clarity or consistency.	Most support is relevant, adequate and logical.  Data, quotation, and information from sources is smoothly incorporated and has adequate comment.  Mostly clear citation, but minor issues with consistency or clarity.	Support is compellingly logical, relevant, and thorough.  Makes thorough, thoughtful, analytic comment on data, quotation, and information from sources,  Citation is meticulously consistent and appropriate for the context and discipline.
COHERENCE logical, clear organization	Poor organization or lack of organization heavily interferes with the message  Layout and presentation is confusing or inappropriate	Organization requires the reader to supply connections.  Layout is simplistic and sometimes confusing	Organization is clear, argument is easy to follow. It rarely requires the reader to supply connections  Layout and presentation is clear and effective	Has a strong and compelling direction and flow, aided by transitions or other markers.  Uses a professional and appropriate layout and presentation
LANGUAGE USE conventions of spelling, punctuation, grammar, word choice, and tone	Pervasive errors significantly interfere with meaning.  The language is consistently too casual or confusing for a professional context.	Includes several errors that interfere with meaning.  The language is often too casual or confusing for an academic or professional context	May include several errors that do not significantly interfere with meaning.  The language is mostly clear, specific and professional.	May include a few minor errors that do not interfere with meaning.  Language is so specific, vivid, appropriately creative, or professional that it adds to the power of the argument.