

2022-2023 Quantitative Literacy Assessment Report

Anna Aboud, ILO Lead Assessment Specialist

*Students will apply relevant scientific, mathematical, and logical methods to analyze and solve problems effectively and be able to utilize the results appropriately when making decisions.
(Westmont Quantitative Literacy ILO and Quantitative and Analytical Reasoning GELO)*

Contents

1	Rationale	2
2	Process	3
3	Results	4
3.1	QR Relevance	5
3.2	QR Quality	5
3.3	Problematic Characteristics	7
4	Conclusion	8
5	Suggestions for Future Assessments	10
6	Acknowledgements	11
	Appendix A	12
	Appendix B	15
	Appendix C	17

1 Rationale

In 2022-2023 Westmont assessed the Quantitative and Analytical Reasoning (QAR) General Education Learning Outcome and Quantitative Literacy (QL) Institutional Learning Outcome using the novel approach of evaluating quantitative literacy through student writing across all disciplines.

¹

The literature identifies four requisite facets of effective Quantitative Reasoning (QR): (1) Basic mathematical and statistical skills; (2) Implementation of these skills in context; (3) Communication of contextualized applications of QR; (4) The habit of mind to approach questions with a quantitative lens. Historically, many QR tests assess only the first of these four facets. However, as described by Lynn Steen in her seminal national report on Quantitative Literacy, “The test of numeracy [quantitative reasoning], as of any literacy is whether a person naturally uses appropriate skills in many different contexts.” The spirit of this quote is captured in facets (2)-(4).^{2,3}

As part of their Quantitative Inquiry, Reasoning, and Knowledge (QuIRK) initiative, Carleton College created and tested an innovative assessment tool to address facets (2)-(4). The initiative was supported by grants from the U.S. Department of Education’s Fund for the Improvement of Post-Secondary Education, the National Science Foundation, and the W.M. Keck Foundation. The QuIRK assessment tool measures the quantity and quality of the QR implementation using student writing samples from across all disciplines.

Westmont utilized a traditional QR multiple-choice test (the Quantitative Literacy Reasoning Assessment test developed by Bowdoin College) for its last QAR/QL assessment cycle in 2015-2016. Although the current QAR/QL assessment team recognizes the value of measuring basic mathematics and statistics skills in this manner, the team also desired to measure student proficiency in facets (2)-(4) of effective QR. The assessment team decided to implement the QuIRK tool in particular (assessing QR using student writing from across disciplines) for the following reasons:

1. Assessing QR through writing naturally supports the liberal arts emphasis on developing well-rounded individuals who think critically and communicate effectively within and across a variety of disciplines.
2. In order to responsibly navigate our increasingly data-driven world, students must expand their facility in the recognition and use of QR as a rhetorical tool. Effective QR enables individuals to build stronger arguments and communicate information more clearly (and ethically!) in both personal and professional spaces (e.g., applying for grants for a nonprofit, making a presentation to a school board about the potential impact of a new program, making the case for a pay raise). Conversely, savvy consumption of material claiming quantitative

¹Note on Terminology: the acronyms QAR and QL refer specifically to the names of Westmont’s General Education Learning Outcome and Institutional Learning Outcome, respectively. Although there is no consensus on terminology at a national level, the phrase Quantitative Reasoning (QR) is used throughout the relevant literature, and we will default to this terminology in our report when speaking of this skill in general.

²Steen, Lynn Arthur, ed. *Mathematics and democracy: The case for quantitative literacy*. Princeton, NJ: NCED, 2001.

³Grawe, Nathan D., and Carol A. Rutz. “Integration with writing programs: A strategy for quantitative reasoning program development.” *Numeracy* 2, no. 2 (2009): 2.

evidence is vital to making informed decisions which impact both the individual and society (e.g., voting, political policy, vaccine decisions, retirement plans). This discerning is particularly salient in our age of widespread misinformation as well as the staggering quantity of information generated by Artificial Intelligence.

3. This approach emphasizes that QL is an Institutional Learning Outcome, not just a skill that is confined to mathematics courses. After conversations with the Carleton QuIRK team, the Westmont assessment team hoped that using the QuIRK tool (as well as the follow-up professional development opportunities) would increase faculty awareness of the value of effective QR for all majors. Just as all faculty members play a role in helping students develop more effective and sophisticated writing skills (in a discipline-appropriate quantity and manner), all faculty members play a role in helping students develop more effective and sophisticated QR skills (in a discipline-appropriate quantity and manner). This idea of collective responsibility and benefit broadens the conversation surrounding the QL Learning Outcome, enabling Westmont to mobilize its most valuable instructional resource (the faculty!) to address student weaknesses in QR.

2 Process

Writing samples were collected from 203 seniors in the spring of 2023. Of these samples, 17 had to be discarded for process/content issues (blank files, incorrect files, unassessable type of submission), leaving 186 assessable samples. Students were asked to submit a writing sample fulfilling the following criteria:

- You are proud of your work on this sample.
- You completed this sample during your final two years at Westmont.
- The sample is (approximately) 750-4000 words long.
- It is preferable that the sample is from a course within your major (or one of your majors). If this is not possible, however, you may choose a sample from any course you have completed at Westmont.
- Preferably (although it is not required!) your sample should reference quantitative information (e.g., numbers!), either as a main component of the paper or as an incidental fact (e.g., 40% of people think...).

Students in many majors were also given a list of examples of one or more sample assignments from specific courses which would fit the above criteria (these lists were provided by faculty members from the major in question). The final bullet point in the criteria above was an attempt to ensure that we received enough samples using QR without overprescribing the sample type. In this way we hoped to catch samples in which students were not using QR, even though it would be appropriate/relevant. The Carleton assessment team reported that this was a key subset of student samples to examine.

Writing sample collection was primarily coordinated through partnerships with senior seminar or capstone course instructors (78% of the faculty teaching these courses agreed to be partners in the assessment). Faculty partners distributed the task instructions to their students, explained the value of the task, and in some cases gave a point value to task completion (not all instructors did this, but it is highly recommended!). During the spring semester instructors were sent periodic

submission updates for students in their course. The (nontrivial number of) seniors not enrolled in senior seminar/capstone courses were (regularly) contacted by the Lead Assessment Specialist via email and also received a follow-up email from their department chair if they had not made a submission by a particular date. Slides about the assessment were also posted in the Dining Commons to help build student awareness of the project.

A team of 13 faculty/staff members from across the institution met on May 15-19, 2023 to assess the student writing samples. On the first day of the project, the reviewer team was trained on the QuIRK rubric as well as the assessment platform (Chalk & Wire). For the next two days, each sample was independently scored by two assessors. Writing samples were categorized as QR Centrally Relevant, QR Peripherally Relevant, or not QR relevant (brief descriptions of these categories are given in Table 1). Any discrepancies in this categorization were resolved by discussion between the two involved assessors. The QR Centrally and Peripherally Relevant samples were then awarded a performance level on a scale of 1 to 4 (level 1=deficient and level 4=excellent) according to the quality of QR integration demonstrated and the effectiveness of the QR in furthering the stated or implied goals of the sample. For a complete description of each performance level, see Appendix A. Discrepancies of more than 1-point value on this 4-point scale were resolved by discussion between the two involved assessors. Each of the two assessor scores was preserved (rather than averaging the values), resulting in n=372 data points from the 186 usable student samples. Assessors then marked each sample for the presence or absence of a list of eight “Problematic Characteristics” (e.g., “Fails to provide numbers that would contextualize the argument.” See Table 2 or Appendix A for a complete list.). On the afternoon of the final day of the assessment, the review team discussed the results, reflected on the assessment experience, and made suggestions for effective interventions and future assessments.

Not QR Relevant	Peripherally QR Relevant	Centrally QR Relevant
No potential uses of numbers or miscellaneous uses only.	Potential uses of numbers to provide useful detail, enrich descriptions, present background, or establish frames of reference.	Potential uses of numbers to address a central question, issue, or theme.

Table 1: QR relevance categories

As part of addressing the QAR General Education Learning Outcome component of this assessment, Westmont’s General Education Committee also completed a syllabus audit of relevant QAR certified courses. The results of this audit are given in Appendix C.

3 Results

We resisted the urge to overanalyze the quantitative data collected in this assessment. The number of usable samples collected (186) was modest, meaning that when the data was disaggregated—particularly according to multiple criteria—the sample sizes became very small. The assessment team still believes that much can be learned about student proficiency in QR by holistically examining the numerical results of the assessment and enriching this examination with feedback and insights from the faculty/staff reviewer team.

3.1 QR Relevance

Of the collected samples, 34% were categorized as not QR Relevant, 23% were categorized as Peripherally QR Relevant, and 42% were categorized as Centrally QR Relevant. Among the Natural and Behavioral Sciences, the breakdown was 20% not QR relevant, 21% QR Peripherally Relevant, and 58% Centrally Relevant. Percentage breakdowns in the Social Sciences and Humanities were less helpful as they were skewed by significant outlier majors with large sample sizes (e.g., Economics and Business with 42 student samples, 55% of which were categorized as not QR Relevant). Most majors in these divisions also demonstrated distinctive patterns which were lost in the aggregate data. Majors with particularly notable patterns were Sociology (80% of samples QR Centrally Relevant) and Political Science (100% of samples QR Centrally Relevant). (**Key Takeaway #1**, **Key Takeaway #5**)

3.2 QR Quality

At this point, samples which were categorized not QR relevant were taken from the pool, leaving 246 QR relevant data points. Among these data points, the percentage of QR scores at each performance level are represented in Figure 1. The mean scores for Overall QR, QR Centrally Relevant, and QR Peripherally Relevant were 2.1, 2.4, and 1.7, respectively. It is worth noting that scores were significantly stronger for the QR Centrally Relevant category than for the QR Peripherally Relevant category ($p < .0001$). This suggests that Westmont students are more effectively using Quantitative Reasoning in areas which more obviously/explicitly call for QR (lab reports, quantitative research projects, etc.), but are less adept at using QR in situations where it is not a central part of the assignment (perhaps because students are not aware that there can still be effective or ineffective integration of QR in these contexts). (**Key Takeaway #2**)

We disaggregated the QR quality (Overall, Central, and Peripheral) results by 1st Generation Status, Ethnicity, Transfer Status, IPEDS, and Gender. With the exception of two demographic categories (both with two or fewer students), performance was poor in all categories (mean scores for Centrally Relevant QR ranged from 2.1-2.5 and mean scores for Peripherally Relevant QR ranged from 1.4-1.8). We found no statistically significant differences between the mean scores in any demographic category; however, the overall trend of higher competency in Central QR than Peripheral QR was followed in every category. To see a full list of the disaggregated scores, we refer the reader to Table 3 in Appendix B. We reiterate that none of the differences between mean scores between demographic groups are statistically significant. We also disaggregated the results

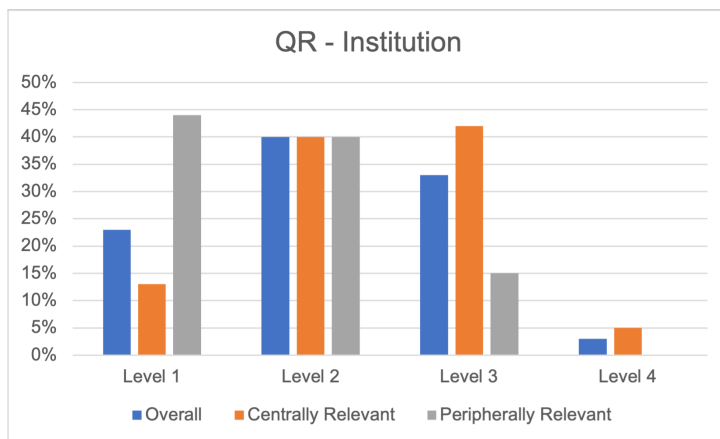


Figure 1: QR institutional scores by performance level

according to Race 1, Race 2, Major 2, and Minor. This data is available upon request but is not included in this report on account of the lack of relevance to the current discussion (due to a variety of factors such as sample size and lack of statistically significant results). (**Key Takeaway #3**)

To give a sense of the types of score distributions one may see across levels, we expanded two of the demographic breakdowns into graphs depicting the percentage of students achieving each performance level within both the Central QR and Peripheral QR categories (Figures 2-5).

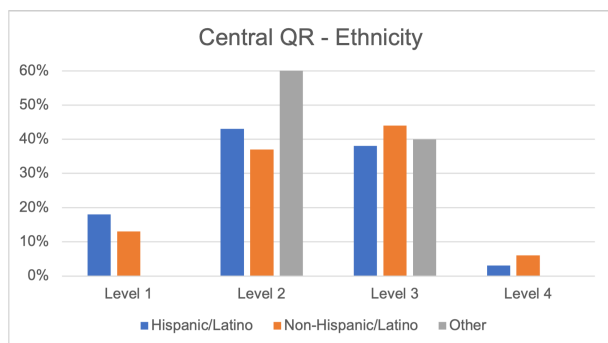


Figure 2: Performance level on central QR by ethnicity

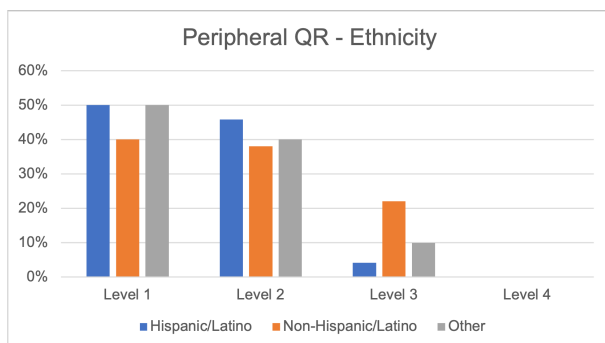


Figure 4: Performance level on peripheral QR by ethnicity

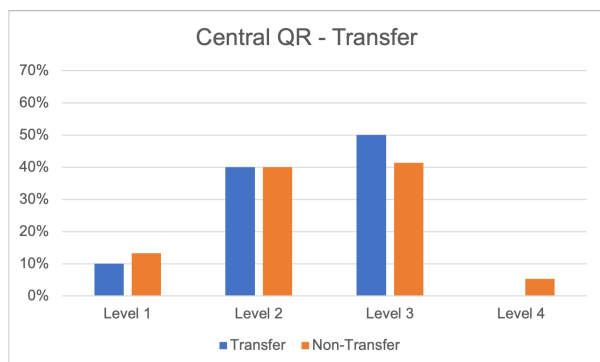


Figure 3: Performance level on central QR by transfer status

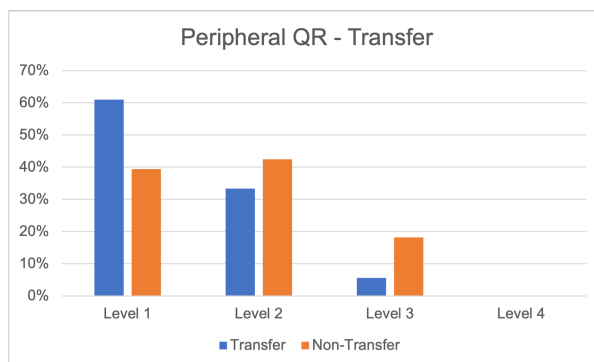


Figure 5: Performance level on peripheral QR by transfer status

As alluded to previously, different majors demonstrated distinctive patterns in assignment type, QR Relevance, and QR Quality. Interesting trends were also noted by the faculty reviewer team. Individual departments are encouraged to contact the Lead Assessment Specialist if they would like a further breakdown of numerical results for their majors as well as an overview of qualitative observations from the faculty reviewer team. QR Quality scores broken down by major are provided in Table 4 in Appendix B. Because of the small sample sizes and distinctive assignment types chosen by students, the assessment team believes that this data is best interpreted in the context of a partnership between the assessment team and the department, where both parties can provide much needed context and nuance to the conversation. (**Key Takeaway #5**)

3.3 Problematic Characteristics

The rubric included a list of eight “Problematic Characteristics” which were marked as present or not present in each of the QR relevant samples. These characteristics provide a helpful breakdown of the most prevalent issues weakening student QR reasoning, and consequently supply concrete foci for intervention efforts on a course and institutional level. Table 2 shows the percentage of QR relevant samples which contained each of these characteristics (a Problematic Characteristic was only marked as present when it detracted from the writer’s argument).

Problematic Characteristic (PC)	QR Centrally Relevant Samples with PC	QR Peripherally Relevant Samples with PC
PC 1: Uses ambiguous words rather than numbers.	71%	68%
PC 2: Fails to provide numbers that would contextualize the argument.	57%	69%
PC 3: Fails to describe own or others data collection methods.	41%	33%
PC 4: Doesn’t evaluate source or methods credibility or limitations.	56%	48%
PC 5: Inadequate scholarship on the origins of quantitative information cited.	30%	28%
PC 6: Makes an unsupported claim about the causal meaning of findings.	42%	33%
PC 7: Presents numbers without comparisons which might give them meaning.	42%	32%
PC 8: Presents numbers but doesn’t weave them into a coherent argument.	20%	17%

Table 2: Percentage of samples containing each problematic characteristic (PC)

Interestingly, the same problematic characteristics appeared most often in both the Peripherally Relevant and Centrally Relevant QR samples (although they manifested themselves differently in the two different contexts). This suggests that targeting interventions based on these three characteristics could have a strong impact on student QR implementation in both areas. (**Key Takeaway #4**)

4 Conclusion

Much of the value of this assessment process was in the rich reflections of the faculty/staff reviewer team. Each reviewer assessed dozens of student papers according to the project rubric, gaining an excellent sense of not just what score students were receiving, but why students were receiving this score (where/how they were falling short!). Reviewer team reflections on the numerical results, the process, and the ILO itself lent color, nuance, and context to the numerical results. As experts within their disciplines, faculty reviewers were able to provide valuable major-specific context for many student samples as well as offer insight on the types of interventions which might be most helpful and impactful within their majors. The key takeaways and action items listed below are largely generated from these conversations.

Key Takeaways

1. QR is relevant for students within all disciplines at Westmont. QR is centrally relevant to thinking and arguing well within the Natural, Behavioral, and Social Sciences. QR is peripherally relevant to all disciplines at times and there are significant negative consequences when students are not able to use QR well. The review team believes all Westmont faculty are interested in helping our students make better arguments and write more clearly. Furthermore, there is a simple path to improvement in this area through low effort and high impact actions.
2. Westmont students are performing poorly in QR. Students performed slightly better in contexts where the need for QR is more apparent (lab reports, quantitative research studies, etc.). They were worse at recognizing and using QR as a rhetorical tool in peripheral contexts. As both Central and Peripheral QR are vital tools for sophisticated and ethical consumption and production of information, this is quite concerning. Students may benefit from clearer expectations in this area both in terms of assignment descriptions and rubric items.
3. Students are performing comparably across all considered demographic groups with no statistically significant differences in mean performance scores. It is worth noting that some demographic groups had no samples.
4. The most common Problematic Characteristics hindering the effective use of QR were: 1) Use of ambiguous words rather than numbers; 2) Failure to provide numbers that would contextualize the argument; and 3) No evaluation of source or methods credibility or limitations. The review team posited that many majors/departments may not have structures in place to specifically target these characteristics. Providing resources and time for developing these structures could be a helpful support for faculty.
5. There were definite trends regarding the types of effective and ineffective QR used within major. To interpret these trends, more context is needed for majors where no faculty member was represented in the reviewer team. Conversations with individual departments are encouraged to 1) provide departments with more individualized and nuanced information about the state of their students in this area and 2) gain information about discipline norms which may lend insight to the assessment process.

The goal of this assessment is to positively impact student outcomes in QR. The assessment team believes this goal will be most effectively accomplished at the faculty level. To achieve this end, the entire faculty must be aware of the deficit in this area, see the value of fostering effective QR within their discipline, and be provided with easily implemented and time-efficient paths for addressing these deficits. Most of the action items listed below contribute to this effort.

Action Items

1. Faculty Meeting Report – In the December 8, 2023, faculty meeting the Lead Assessment Specialist will remind faculty of the assessment structure (and the rationale behind it) and share student performance on the assessment.
2. Faculty Forum Workshop – In the January 11 faculty forum, the Lead Assessment Specialist will help faculty see the relevance of this skill for their students by presenting writing samples within a few different disciplines utilizing progressively more advanced levels of QR. Faculty will then be given time and structure to jumpstart their efforts in this area by making a list of assignments from their classes this and next semester where students may find QR helpful.
3. Faculty Professional Development Workshop – In Spring 2024 or Summer 2024 the assessment team hopes to partner with the Writing Center to run a workshop helping faculty to: 1) develop and integrate discipline-appropriate QR prompt paragraphs into their assignment descriptions 2) create QR rubric line items that fit into their preexisting rubrics. After the workshop the team plans to share these resources with the wider faculty, creating a repository of discipline-specific QR prompts and rubric line items. These items would particularly target the QR problematic characteristics identified in this assessment.
4. Departmental Discussions – The Lead Assessment Specialist extends an invitation to each department on campus to discuss the results of their assessment. The assessment team can provide major-specific numerical results, share insights gathered by the faculty reviewer team, and open a conversation about what effective faculty supports in this area may look like.
5. Collaborations with the Writing Center and Library – Over the next year the assessment team plans to have multiple conversations with the Writing Center and Library Staff concerning possible synergistic efforts. One initial idea is to provide short training sessions and resources to Writing Center Tutors in this area. Another is to develop short class presentations that faculty could request for their courses (perhaps integrated with the presentations given by the Library Staff which support the Information Literacy ILO). Both ideas would focus interventions on the top issues identified in the Problematic Characteristics list.

5 Suggestions for Future Assessments

1. Broaden staff/faculty reviewer teams to include assessors from more departments. During this assessment we particularly felt the lack of faculty reviewers from Sociology, Psychology, Political Science, and Economics and Business. Accomplishing this goal may require increasing monetary or time incentives for faculty to participate. For example, it has been noted that the cost of childcare exceeds the amount paid to faculty for their participation. *
2. Broaden staff/faculty reviewer teams to include a representative of the Executive Team and/or a representative of the Board of Trustees. ILO assessment provides a valuable opportunity to understand how students are performing in one of the eight areas that Westmont has deemed vital to our educational program, and we feel that the institution would benefit at multiple levels from the participation of these parties. *
3. Obtaining enough student submissions for any assessment can be a challenge. The most effective method to ensure student submission is to work with course instructors. When instructors give a point value to the submission task, submission rates increase. When instructors set aside class time to give students a chance to complete the task during class, submission rates also increase. *
4. Give more specific guidelines to students about what type of submission to make for this assessment, even if it means running the risk of overprescribing. Some majors submitted a disproportionate number of assignments not relevant to this assessment (for example, many Economics and Business majors submitted a Leadership Essay assignment, none of which ended up being QR relevant). **
5. Although the tool used in this assessment has not proved reliable between institutions (thus eliminating the possibility of the comparison of our students with some national norm), the tool has been shown to be highly reliable within institution. Because of this, the assessment team believes it could be informative to use the same tool in the next QAR/QL assessment cycle. Should capacity allow, the assessment team suggests that the next QAR/QL assessment cycle consist of both writing sample collection (QR in practice!) and also a traditional basic math skills test (such as the Quantitative Literacy Reasoning Assessment test used in 2015-2016). This test would give us a benchmark for national comparison as well as the ability to examine the performance level connections between basic QR skills and the implementation of QR in context. **

* Suggestion for all future ILO assessments

** Suggestion for future QAR/QL assessments

6 Acknowledgements

The Lead Assessment Specialist would like to acknowledge the valuable collaboration and contribution of the following individuals in this assessment project.

- Nathan Grawe, Professor of Economics (Carleton College)
- Tatiana Nazarenko, Dean of Curriculum and Educational Effectiveness (Westmont College)
- Westmont QAR/QL Assessment Team: Carolyn Mitten and Michael Everest
- Westmont Faculty/Staff Reviewer Team: Steve Contakes, Theresa Covich, Stephanie Cowell, Isaac Gomez, Brandon Haines, Bob Haring-Kaye, Ruth Lin, Tim Loomer, Carolyn Mitten, Andrew Mullen, Tatiana Nazarenko, Diane Ziliotto

Appendix A

Following is the QuIRK Assessment Rubric reformatted to match the Chalk & Wire assessment platform.

QR Rubric 2023 Westmont QL ILO Assessment

1. QR Requested

Does the assignment description (if provided) explicitly call for the use of QR in the paper?

NO	YES	No Assignment Description Present	DO NOT SELECT	DO NOT SELECT
----	-----	-----------------------------------	---------------	---------------

2. QR Relevance

Is QR potentially relevant to this paper? This is a reader's assessment of the potential contribution of quantitative information to the paper based on the stated and implied goals of the paper itself; it is not an assessment of the specifications of the assignment. In making this assessment, consider how a reasonable person would consider the relevance of QR to the topic chosen by the student. That is, ask if you would expect QR to play a peripheral or central role in a strong paper on this topic, not if you could somehow squeeze QR into this context.

NO or incidentally only	YES, but peripherally only	YES, centrally	DO NOT SELECT	DO NOT SELECT
No potential uses of numbers or miscellaneous uses only	Potential uses of numbers to provide useful detail, enrich descriptions, present background, or establish frames of reference.	Potential uses of numbers to address a central question, issue, or theme		

3. QR Extent

What is the extent of numerical evidence and quantitative reasoning present in the paper? This is not a rating of the quality of the QR shown; it is an assessment of the degree to which explicit numerical information or reasoning with quantitative information is present.

None	Minimal	Throughout	DO NOT SELECT	DO NOT SELECT
No explicit numerical evidence or reasoning with quantitative reasoning. May include quasi-numeric references (ie "many," "few," "most," "increased," "fell," etc.).	One or two instances of explicit numerical evidence or quantitative reasoning (perhaps in the introduction to set the context), but no more.	Explicit numerical evidence or quantitative reasoning is used throughout the paper.		

4. QR Overall Quality - Centrally Relevant

Level 1	Level 2	Level 3	Level 4	
<p>Use of numerical evidence is so poor that either it is impossible to evaluate the argument with the information presented or the argument is clearly fallacious. Perhaps key aspects of data collection methods are missing or critical aspects of data source credibility are left unexplored. The argument may exhibit glaring misinterpretation (for instance, deep confusion of correlation and causation). Numbers may be presented, but are not woven into the argument.</p>	<p>The use of numerical evidence is sufficient to allow the reader to follow the argument. But there may be times when information is missing or misused. Perhaps the use of numerical evidence itself is uneven. Or the data are presented effectively, but a lack of discussion of source credibility or methods makes a full evaluation of the argument impossible. Misinterpretations such as the confusion of correlation and causation may appear, but not in a way that fundamentally undermines the entire argument.</p>	<p>The use of numerical evidence is good throughout the argument. Only occasionally (and never in a manner that substantially undermines the credibility of the argument) does the paper fail to explore source credibility or explain methods when needed. While there may be small, nuanced errors in the interpretation, the use of numerical evidence is generally sound. However, the paper may not explore all possible aspects of that evidence.</p>	<p>The use of numerical evidence is consistently of the highest quality. When appropriate, source credibility is fully explored and methods are completely explained. Interpretation of the numerical evidence is complete, considering all available information. There are no errors such as confusion of correlation and causation. This paper would be an excellent choice as an example of effective central QR to be shared with students and faculty.</p>	<p>NA Not QR Centrally Relevant</p>

5. QR Overall Quality - Peripherally Relevant

Level 1	Level 2	Level 3	Level 4	
<p>Fails to use any explicit numerical evidence to provide context. The paper is weaker as a result. This paper shows no attempt to employ peripheral QR.</p>	<p>Uses numerical evidence to provide context in some places, but not in others. The missing context weakens the overall paper. Or the paper may consistently provide data to frame the argument, but fail to put that data in context by citing other numbers for comparison. Ultimately, the attempt at peripheral use of QR does not achieve its goal.</p>	<p>The paper consistently provides numerical evidence to contextualize the argument when appropriate. Moreover, numbers are presented with comparisons (when needed) to give them meaning. However, there may be times when a better number could have been chosen or more could have been done with a given figure. In total, the peripheral use of QR effectively frames or motivates the argument.</p>	<p>Throughout the paper, numerical evidence is used to frame the argument in an insightful and effective way. When needed, comparisons are provided to put numbers in context. This paper would be an excellent choice as an example of effective peripheral QR to be shared with students and faculty.</p>	<p>NA Not QR Peripherally Relevant</p>

6. Problematic Characteristic 1

Uses ambiguous words rather than numbers.

Not Present	Present	DO NOT SELECT	DO NOT SELECT	DO NOT SELECT
-------------	---------	---------------	---------------	---------------

7. Problematic Characteristic 2

Fails to provide numbers that would contextualize the argument.

Not Present	Present	DO NOT SELECT	DO NOT SELECT	DO NOT SELECT
-------------	---------	---------------	---------------	---------------

8. Problematic Characteristic 3

Fails to describe own or others' data collection methods.

Not Present	Present	DO NOT SELECT	DO NOT SELECT	DO NOT SELECT
-------------	---------	---------------	---------------	---------------

9. Problematic Characteristic 4

Doesn't evaluate source or methods credibility and limitations.

Not Present	Present	DO NOT SELECT	DO NOT SELECT	DO NOT SELECT
-------------	---------	---------------	---------------	---------------

10. Problematic Characteristic 5

Inadequate scholarship on the origins of quantitative information cited.

Not Present	Present	DO NOT SELECT	DO NOT SELECT	DO NOT SELECT
-------------	---------	---------------	---------------	---------------

11. Problematic Characteristic 6

Makes an unsupported claim about the causal meaning of findings.

Not Present	Present	DO NOT SELECT	DO NOT SELECT	DO NOT SELECT
-------------	---------	---------------	---------------	---------------

12. Problematic Characteristic 7

Presents numbers without comparisons that might give them meaning.

Not Present	Present	DO NOT SELECT	DO NOT SELECT	DO NOT SELECT
-------------	---------	---------------	---------------	---------------

13. Problematic Characteristic 8

Presents numbers but doesn't weave them into a coherent argument.

Not Present	Present	DO NOT SELECT	DO NOT SELECT	DO NOT SELECT
-------------	---------	---------------	---------------	---------------

Appendix B

Table 3 shows the Overall QR, Central QR, and Peripheral QR scores disaggregated by 1st generation status, transfer status, ethnicity, IPEDS, and gender. Table 4 shows Overall QR scores by performance level disaggregated by major. Majors not listed had no QR relevant samples. This could be because the major had no samples submitted or because the samples submitted were not QR relevant. Recall that a sample size of $n = 6$ reflects the work of three students with two scores each. Further note that only QR relevant samples were considered in both tables. As stated in the text body, none of the demographic differences in Table 3 were calculated to be statistically significant.

	Overall	Centrally Relevant	Peripherally Relevant
Overall (n = 246)	2.1	2.4 (n = 162)	1.7 (n = 84)
1st Generation ($n = 38$)	2.18	2.50	1.64
Non-1st Generation ($n = 208$)	2.15	2.37	1.73
Transfer ($n = 28$)	1.79	2.40	1.44
Non-Transfer ($n = 218$)	2.20	2.39	1.79
Hispanic/Latino ($n = 64$)	1.98	2.25	1.54
Non-Hispanic/Latino ($n = 162$)	2.25	2.44	1.82
Other ($n = 20$)	2.00	2.40	1.60
Asian ($n = 12$)	2.25	2.50	1.00
Black/African American ($n = 0$)	—	—	—
Hawaiian/Pacific Islander ($n = 4$)	2.75	2.75	—
Hispanic/Latino ($n = 64$)	1.98	2.25	1.54
White ($n = 126$)	2.27	2.45	1.9
Two or More Races ($n = 16$)	2.07	2.50	1.75
Unknown ($n = 22$)	2.05	2.13	1.83
Non-Resident Alien ($n = 0$)	—	—	—
American/Alaska Native ($n = 2$)	3.00	3.00	—
Male ($n = 100$)	2.10	2.41	1.69
Female ($n = 146$)	2.16	2.37	1.74

Table 3: Mean QR quality scores disaggregated by 1st generation status, transfer status, ethnicity, IPEDS, and gender

Major	Level 1	Level 2	Level 3	Level 4
Art ($n = 6$)	33%	33%	17%	17%
Biology ($n = 28$)	18%	46%	36%	0%
Chemistry ($n = 4$)	100%	0%	0%	0%
Communication ($n = 14$)	29%	50%	21%	0%
Computer Science ($n = 10$)	20%	40%	40%	0%
Data Analytics ($n = 10$)	30%	70%	0%	0%
Economics and Business ($n = 38$)	42%	37%	18%	0%
Engineering Physics ($n = 2$)	0%	50%	50%	0%
English ($n = 6$)	33%	33%	17%	17%
Kinesiology ($n = 44$)	18%	32%	45%	45%
Mathematics ($n = 4$)	50%	50%	25%	25%
Philosophy ($n = 2$)	0%	100%	0%	0%
Political Science ($n = 10$)	10%	40%	50%	0%
Psychology ($n = 22$)	9%	36%	55%	0%
Sociology ($n = 30$)	17%	30%	47%	7%
Spanish ($n = 4$)	0%	100%	0%	0%
Liberal Studies ($n = 8$)	38%	63%	0%	0%
Environmental Science ($n = 2$)	0%	100%	0%	0%
Mechanical Engineering ($n = 2$)	100%	0%	0%	0%

Table 4: Overall QR scores by major given by percentages at each performance level

Appendix C

As a part of the QAR (General Education Learning Outcome) portion of the assessment project, the Westmont General Education (GE) Committee conducted a syllabus audit in Fall 2023. Specifically, the committee examined syllabi for the following courses: CHM-004, CHM-005 (two sections), CHM-005H, MA-005 (two sections), MA-009, MA-010, CS-015, PHS-007, PHS-011, PHY-007, PHY-011, PHY-021, and POL-040 using the established audit templates. The audit found that four Physics syllabi and one Chemistry syllabus required revisions. Dr. Felicia Song, GE Committee Chair, reached out to the respective department chairs regarding necessary revisions. On October 31, 2023, Dr. Nazarenko met with the Physics faculty to update them regarding the GE syllabus requirements.