



Cardinal Core Assessment of Natural Sciences (Fall 2020 Sample)

Cardinal Core Program

The Cardinal Core program at the University of Louisville prepares students to do the advanced work needed for their baccalaureate degrees and prepares them to contribute to society throughout their lives through their professional work and civic engagement. The program emphasizes the development of key intellectual skills relevant to any career path: critical thinking, quantitative reasoning, effective communication, and the understanding of historical, social, and cultural diversity. Students will develop these intellectual skills in the following content areas of Arts and Humanities, Historical Perspectives, Oral Communication, Quantitative Reasoning, Social and Behavioral Sciences, Natural Sciences, Written Communication, and the competency area of Diversity in the United States and Globally. Upon completion of the program, students will be prepared to analyze complex problems and evaluate possible courses of action in an environment characterized by diversity and the need for sustainable solutions.

Assessment Administration

The assessment of student learning outcomes is a national expectation in higher education. Section 8.2.b of the Southern Association of Colleges and Schools Commission on Colleges' (SACSCOC) accreditation standards requires that the institution identify student learning outcomes for collegiate-level general education competencies in its undergraduate degree programs, assess the extent to which it achieves these outcomes, and provide evidence of seeking improvement based on analysis of the results. Further, the Kentucky Council on Postsecondary Education (CPE) states that "All Kentucky public universities and KCTCS colleges are expected to assess, in accordance with SACSCOC Principles of Accreditation and based upon nationally accepted standards, the student learning outcomes associated with their general education programs, indicate a relationship to the faculty-generated Statewide General Education Student Learning Outcomes, and provide evidence of ongoing assessment that ensures comparability for transfer purposes on a three-year cycle."

The Cardinal Core Curriculum Committee (CCCC) is charged with continued oversight of the assessment of student learning outcomes across the Cardinal Core curriculum to support the continuous improvement of the Cardinal Core program in alignment with SACSCOC and CPE requirements. The assessment operates on a three-year cycle, in which samples of student work are collected from one content area each semester and assessed by a panel of trained faculty. The CCCC began a pilot of the Association of American Colleges and Universities (AAC&U) VALUE (Valid Assessment of Learning in Undergraduate Education) Rubrics for Critical Thinking, Written Communication, Oral Communication, Intercultural Knowledge and Competence, and Quantitative Literacy for assessments of the overarching intellectual skills of critical thinking, effective communication, quantitative reasoning, and social, historical, and cultural diversity in the Arts and Humanities, Historical Perspectives, Oral Communication,

Quantitative Reasoning, Social and Behavioral Sciences, and Written Communication content areas. The CCCC is also piloting a new University of Louisville Natural Science Critical Thinking Rubric for the assessment of critical thinking skills in the Natural Sciences.

The Fall 2020 assessment was focused on courses in the Natural Sciences (S, SL, and B) content area. Samples of student work from S, SL, and B courses were assessed using the new Cardinal Core Natural Sciences Critical Thinking Rubric. The University of Louisville Outcomes, Kentucky Statewide Outcomes, and the AAC&U VALUE Rubric Measures used to assess courses in the Natural Sciences content area are provided below.

University of Louisville Natural Sciences Learning Outcomes

Natural Sciences (S, SL, and B) are concerned with understanding the laws of nature and the physical world. Students who satisfy this requirement will be able to do all of the following:

1. Demonstrate an understanding of the nature and methods of science inquiry.
2. Apply scientific principles: to interpret evidence, to make predictions, and/or to explain cross-cutting concepts in one or more of the sciences.
3. Explain how scientific principles relate to issues of personal and/or societal importance.
4. Communicate effectively an understanding of scientific concepts and experimental outcomes in speech or writing, using sound scientific terminology and citation appropriate to the discipline.

Statewide Natural Sciences Student Learning Outcomes

1. Demonstrate an understanding of the methods of science inquiry.
2. Explain basic concepts and principles in one or more of the sciences.
3. Apply scientific principles to interpret and make predications in one or more of the sciences.
4. Explain how scientific principles relate to issues of personal and/or public importance.

University of Louisville Natural Sciences Critical Thinking Rubric

- (NS1) Demonstrate an Understanding of Methods of Science
- (NS2) Apply Scientific Principles
- (NS3) Connecting Scientific Principles to Issues of Personal and/or Societal Importance
- (NS4) Communicate an Understanding of Scientific Concepts and Experimental Outcomes

The Natural Sciences Rubric uses the same scoring categories as the AAC&U VALUE Rubrics to ensure consistency across all instruments used in the Cardinal Core Assessment, with 4 indicating performance of the measure as “capstone” level, 3 indicating performance at “milestone,” 2 indicating “milestone,” and 1 indicating performance at “benchmark.” In addition, a score of zero can be assigned to any work that does not meet the benchmark level performance. The University of Louisville further disaggregates the zero option into a “not applicable” rating that can be selected for assignments that did not provide an opportunity for the student to

demonstrate the criterion within the rubric measure, as opposed to the student simply not meeting the rubric criteria.

Assessment Process

The Natural Sciences (S, SL, and B) Assessment was originally planned for Spring 2020, however due to the COVID pandemic, the assessment was postponed until Fall 2020 to allow faculty the opportunity to adapt to new teaching methods and ensure a sample that more accurately represents a traditional semester at the university. In preparation for the assessment, the Cardinal Core Office collaborated with the department chairs of departments that offer Natural Science courses regarding the details of the upcoming assessment to ensure faculty participation and appropriate sampling. A formal memo outlining the project and process was provided to all faculty teaching S, SL, and B Cardinal Core courses prior to the start of the semester to ensure a mutual understanding of project expectations. The initial communication provided a timeline for collection of assignment prompts and student work.

Mid-semester, the Cardinal Core Office retrieved the class rosters for all S, SL, and B Cardinal Core courses from the Office of the Registrar and selected a stratified random sampling, to ensure that the sample included students from all courses. In the instance of a faculty member teaching multiple sections of the same course, not all sections were included in the random sample to limit the number of samples coming from one faculty member. Instructors of all S, SL, and B courses were sent a list of students selected for the assessment along with detailed instructions requesting that instructors provide a copy of one assignment prompt along with the ungraded responses for the selected students to be sent via email to the Cardinal Core Office service account.

Student artifacts were collected and stored in an electronic repository and uploaded into the LiveText© assessment management system. A panel of faculty (tenured and tenure-track faculty, term faculty, and adjunct faculty) and graduate teaching assistants assessed student artifacts. The University of Louisville Natural Sciences Critical Thinking Rubric was applied to all student artifacts.

Prior to the assessment reading, assessors completed mandatory training/rubric norming. Due to the COVID pandemic, and in contrast to previous assessments, the training was conducted completely online in Blackboard. Training materials were developed in collaboration with the CCCC Assessment Subcommittee and based upon UofL's long-standing general education assessment training practices, as well as AAC&U VALUE Institute training procedures. The training module consisted of an overview of the assessment process and holistic assessment practices, a dissection of the Natural Science rubric, and scoring of benchmark sample assignments. Benchmarks were assignments selected to represent a wide range of content and skill development in order to give the assessors a baseline for measuring expectations of learning and evaluating student performance (Herman, Osmundson, & Dietel, 2010). Assessors were given a week to complete the training module in Blackboard and submit their scores for all benchmark samples. The results of the benchmark scoring were compiled and sent out to all assessors. Competency areas (rubric rows) with disagreement among reviewers were discussed in the findings shared with reviewers to clarify intended applications of the rubrics. The results

from scoring benchmark samples for the Natural Sciences Critical Thinking Rubric is provided in Appendix A.

After completion of the assessment training/rubric norming, each assessor was assigned a username and password for one of three LiveText© accounts and a list of courses and sections to assess. Three readers assessed each artifact so that scores could be compared across assessors for reliability purposes. Assessors were given a week to complete all assessments.

Data Collection Overview

The enrollment for Natural Sciences (S, SL, and B) Cardinal Core courses in Fall 2020 was approximately 7900 students at the time the sample for the assessment was selected. (The Cardinal Core Office has traditionally waited to identify a sample after the deadline to withdraw from courses has passed, but that deadline was extended in Fall 2020 and would have been too late in the process to notify instructors of the identified students. As a result, the Cardinal Core Office requested a larger sample to account for students that may have withdrawn after the sample was identified.) The Cardinal Core Office collected a stratified random sample and received 674 student work samples. The sample included 32 work samples from Anthropology, 294 from Biology, 148 from Chemistry, 51 from Geography and Geosciences, 25 from Public Health, and 124 from Physics. 398 work samples came from lecture (S) courses, 256 work samples came from lab (SL) courses, and 20 work samples came from combine lecture/lab courses (B).

Summary of Assessment Data

For the assessment of Natural Sciences (S, SL, and B) outcomes, 674 student artifacts were assessed by faculty and graduate teaching assistants using the University of Louisville Natural Sciences Critical Thinking Rubric.

Tables 1-3 provide the percentage of work samples scored at each rubric level for the Natural Sciences Critical Thinking Rubric. Assessors were given the option for a “0” score and for purposes of understanding why a “0” was assigned, the Cardinal Core Assessment provided reviewers with the option of “Not Applicable”. The “Not Applicable” indicates an absence of the assessment criteria due to the type of assignment, while a “0” indicates that the student could have demonstrated the criteria and did not.

Table 1

Percentage of Artifacts Scored at Each Rubric Level for Natural Sciences

	NS1	NS2	NS3	NS4
Capstone (4)	21.1%	16.5%	9.7%	12.0%
Milestone (3)	32.4%	28.9%	15.0%	24.0%
Milestone (2)	29.1%	21.1%	16.0%	28.9%
Benchmark (1)	10.2%	7.7%	8.4%	9.3%
(0)	2.1%	1.7%	4.5%	1.2%
Not Applicable (NA)	5.0%	24.2%	46.3%	24.6%

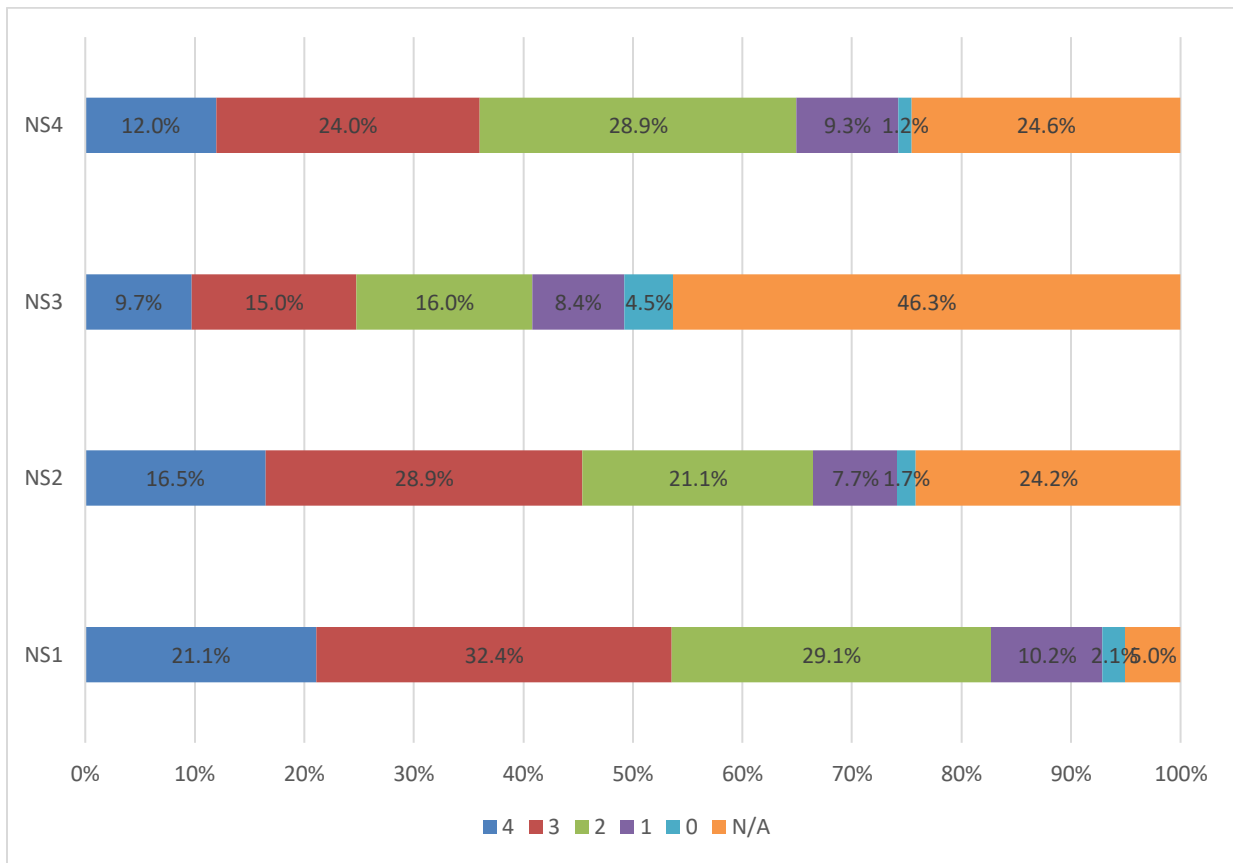


Figure 1. Percentage of Artifacts Scored at Each Rubric Level for Natural Sciences

Table 2 and Figure 2 provide a breakdown of the N/A ratings by S, SL, and B courses.

Table 2

Percentage of Not Applicable Scores by S, SL, and B Courses

	NS1	NS2	NS3	NS4
S	66.7%	74.4%	46.6%	91.1%
SL	32.4%	24.9%	49.3%	8.7%
B	1.0%	0.6%	4.1%	0.2%

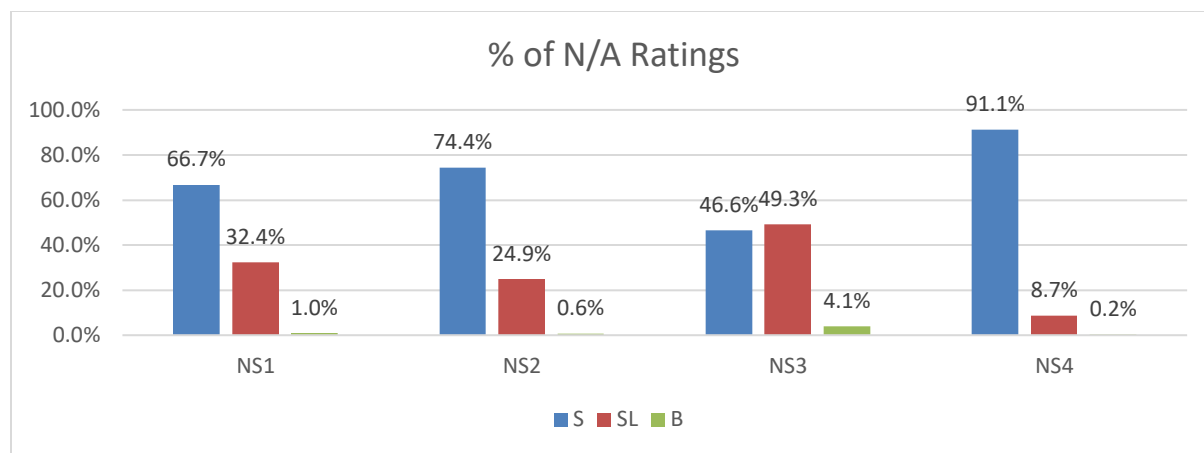


Figure 2. Percentage of Not Applicable Scores by S, SL, and B Courses

In alignment with AAC&U VALUE Institute practices, scores for each individual work sample were calculated based upon scores assigned by three separate reviewers. The scores from all three reviewers were averaged and rounded to determine individual work sample scores for each rubric row. The mode for the individual work sample scores is provided in Table 3.

Table 3

Mode of Individual Work Sample Scores

	Mode
NS1 – Demonstrate an Understanding of Methods of Science	3
NS2 – Apply Scientific Principles	2
NS3 – Connecting Scientific Principles to Issues of Personal and/or Societal Importance	0 or N/A
NS4 – Communicate an Understanding of Scientific Concepts and Experimental Outcomes	2

Inter-rater Reliability

Three separate readers assessed each student artifact. Table 4 displays the mean score for the three separate readings of all artifacts.

Table 4

Inter-rater Summary for University of Louisville Natural Sciences Critical Thinking Rubric

	Assessor 1	Assessor 2	Assessor 3	Standard Deviation
NS1	2.50	2.30	2.71	0.20
NS2	2.16	1.47	2.44	0.50
NS3	1.45	1.11	1.17	0.18
NS4	2.20	1.59	1.82	0.31

In addition to the descriptive statistics, Table 5 provides multiple measures of inter-rater reliability. The percentage agreement value was calculated to determine the percentage of artifacts for which all three assessors scored at the either the same or within one performance level. Values for *Total Agreement* provided in Table 5 represent the percentage of artifacts for which all three assessors selected the same score (e.g., Assessors 1, 2, and 3 all selected 3). *Agreement (within 1 level)* represents the percentage of artifacts for which all three assessors scored the artifact at the same performance level or within one level (e.g., Assessor 1 selected a score of 3, Assessor 2 selected a score of 2, and Assessor 3 also selected a score of 2). If the assessor assigned “not requested” for the artifact that was treated as a 0 for the inter-rater reliability analysis since a 0 and “not requested” would both indicate the reviewer did not see the student demonstrate any component of the rubric measure.

In addition to percentage agreement, a one-way, average-measures intra-class correlation coefficient (ICC) was calculated to assess inter-rater reliability. ICC coefficients between .75 and 1.00 are considered excellent, .60 to .74 considered good, .40 to .59 fair, and below .4 is considered poor (Cicchetti, 1994).

Table 5

Inter-rater Reliability for University of Louisville Natural Sciences Critical Thinking Rubric

Competency Measure	Total Agreement	Agreement (within 1 level)	ICC	95% Confidence Interval
NS1	16.2%	59.9%	0.56	.502-.617
NS2	13.1%	37.7%	0.44	.365-.511
NS3	38.3%	60.8%	0.76	.723-.787
NS4	16.6%	52.4%	0.69	.641-.724

Syllabus Review

The Provost requests that all faculty load their syllabi to Blackboard© each semester. These syllabi are then available through the university’s course catalog system. For the purpose of this review, the Cardinal Core Office collected all Natural Sciences (S, SL, and B) syllabi that were loaded to Blackboard in Fall 2020.

The review of syllabi sought to answer two questions:

- 1) Does the syllabus contain the content specific Cardinal Core learning outcomes approved for the course?
- 2) Are assessment methods stated that support the content-specific Cardinal Core learning outcomes approved for the course?

The syllabus review included syllabi from 148 S, SL, and B Cardinal Core course sections offered in the Fall of 2020. The review of 148 syllabi identified 117 syllabi (79.1%) containing the Natural Science Cardinal Core learning outcomes approved for the course. Further review of

the syllabi revealed that 107 syllabi (72.3%) also listed the assessment methods for the Cardinal Core outcomes.

The Cardinal Core Curriculum Committee (CCCC) has continued to emphasize the importance of incorporating the Cardinal Core learning outcomes into course syllabi. Integration of the Cardinal Core outcomes into the syllabus is one indication to the committee that faculty are incorporating the learning outcomes into the course curriculum.

Summary and Plan for Improvement

The results of this assessment are intended to serve as a baseline for future assessments in the Natural Sciences content area and intended to help guide continuous program improvement. Of the four Natural Science competencies, students performed best on “demonstrating an understanding of the methods of science” and did not perform as well on “applying scientific principles”, “connecting scientific principles to issues of personal and/or societal importance”, and “communicating an understanding of scientific concepts and experimental outcomes”. As noted in the results provided in Table 1, nearly half of all work samples from the S, SL, and B courses did not ask students to make connections between scientific principles and issues of personal and/or societal importance. Departments are asked to revisit the outcomes to ensure all outcomes are being addressed and assessed within their courses. Also, faculty are asked to consider the assignments that they select for future assessments, to make sure that the student work they select demonstrates all outcomes.

This was the first Natural Sciences assessment that included assignments from large lecture (S) courses that only use multiple-choice assignments. It was expected that assignments from these large lecture courses would receive a score of “N/A” for NS4 (Communicate an Understanding of Scientific Concepts and Experimental Outcomes) due to the lack of written work. Many of the large lecture courses approved for Cardinal Core were approved with the caveat that outcome 4 (NS4) would be met in the corresponding lab course. As shown in Table 2, science lecture (S) courses account for 91.1% of the N/A ratings for NS4, with most of the remaining N/A ratings coming from science lab courses (SL). Although S courses make up approximately 60% of the assessment sample it is important to note that work samples from S courses make up a large percentage of the N/A ratings for all Natural Science outcomes. As a result of these findings, the CCCC will help to identify some examples of multiple-choice questions/assignments that provided assessors the opportunity to assign a score and share these samples with departments in preparation for future assessments, to help guide the selection of work samples. The CCCC continues to encourage more small seminar type classes that provide students opportunity to demonstrate more active engaged learning and communication of scientific concepts, as originally proposed by the General Education Task Force when developing the Cardinal Core program.

After transitioning some components of the assessment training to online in spring 2020, the CCCC recommended shifting more of the assessment training to online to align with AAC&U

VALUE Institute practices and to improve inter-rater reliability. As a result of this recommendation and the COVID pandemic, the entire training process was transitioned to online. Inter-rater reliability was good for all Natural Science competencies, except for NS2 (Applying Scientific Principles). Based on informal feedback from assessors, there was confusion around what is considered making “predictions about natural phenomena” as noted in the performance categories. This will be further clarified in context of the different disciplines for future assessments. Based upon observations of the Director of Cardinal Core Assessment and feedback from assessors, it is recommended that the majority of the training process continue online, however bring back a synchronous (face-to-face or online) meeting opportunity to discuss the results of the practice scoring and discuss any areas of disagreement among reviewers. The Cardinal Core Office conducted an informal follow-up survey for the Natural Sciences Assessment and approximately 70% of respondents indicated a synchronous (online or virtual) wrap-up session for training would be beneficial.

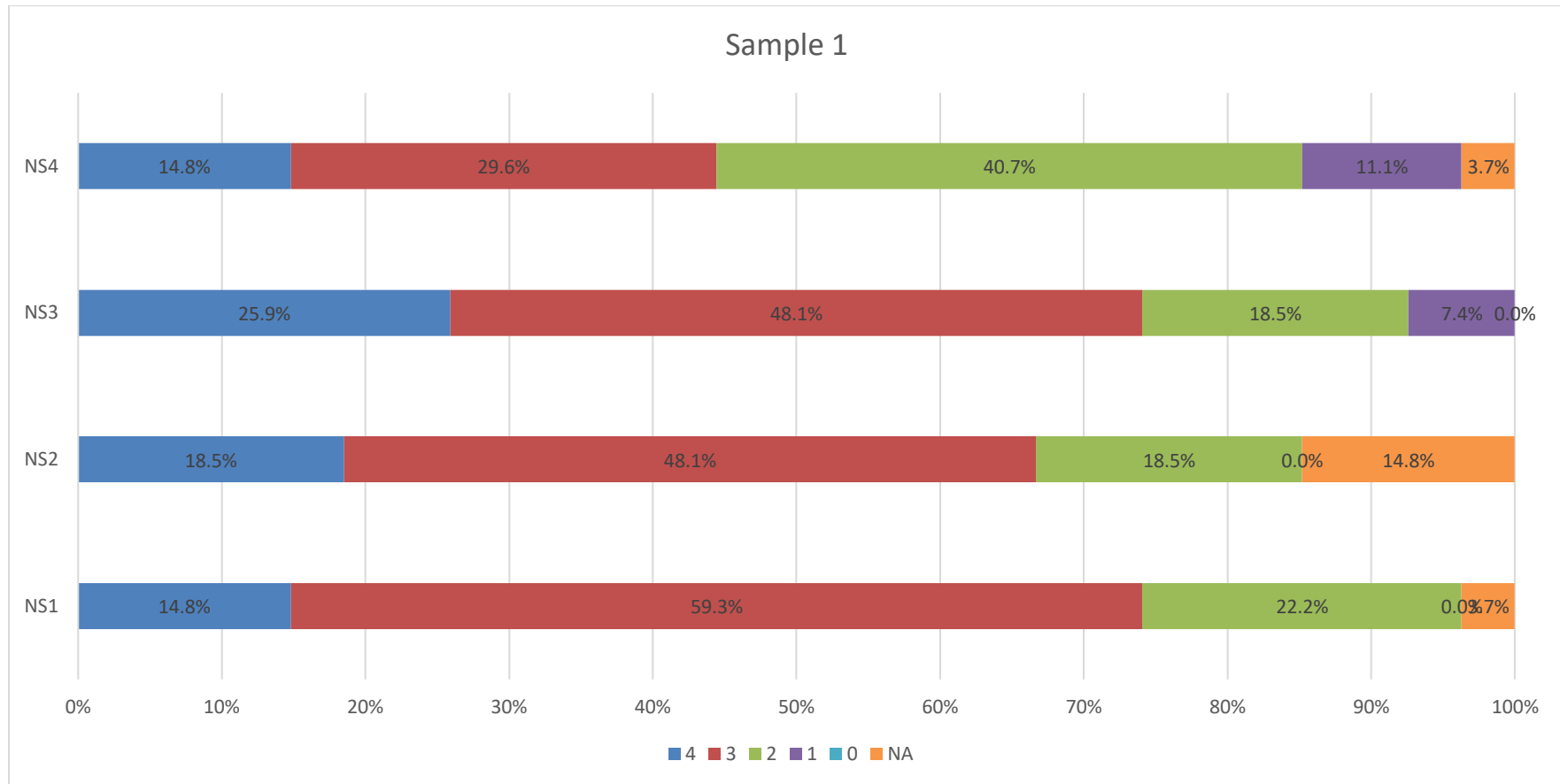
In conclusion, the results of this assessment are intended to help faculty identify potential strengths and areas for growth in the Cardinal Core curriculum. While this first cycle of assessment has some limitations, the data indicate students are performing better on some measures or outcomes than others and that work samples identified for the assessment need to more closely align with the Natural Science outcomes in order to measure student performance. It is recommended that faculty take this into consideration in their own curriculum and assignments to ensure that students are developing the knowledge, skills, and dispositions outlined in the Cardinal Core program.

References

- Association of American Colleges and Universities. (2009). *Critical Thinking VALUE Rubric*. Retrieved from <https://www.aacu.org/value/rubrics/critical-thinking>
- Association of American Colleges and Universities. (2009). *Intercultural Knowledge and Competence VALUE Rubric*. Retrieved from <https://www.aacu.org/value/rubrics/intercultural-knowledge>
- Association of American Colleges and Universities. (2009). *Written Communication VALUE Rubric*. Retrieved from <https://www.aacu.org/value/rubrics/written-communication>
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4), 284-290. doi: 10.1037/1040-3590.6.4.284
- Fletcher, I., Mazzi, M., & Nuebling, M. (2011). When coders are reliable: The application of three measures to assess inter-rater reliability/agreement with doctor–patient communication data coded with the VR-CoDES. *Patient Education and Counseling*, 82(3), 341-345. doi:10.1016/j.pec.2011.01.004
- Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology*, 8(1), 23-34.
- Herman, J. L., Osmundson, E., & Dietel, R. (2010). *Benchmark Assessment for Improved Learning*. (AACC Report). Los Angeles, CA: University of California

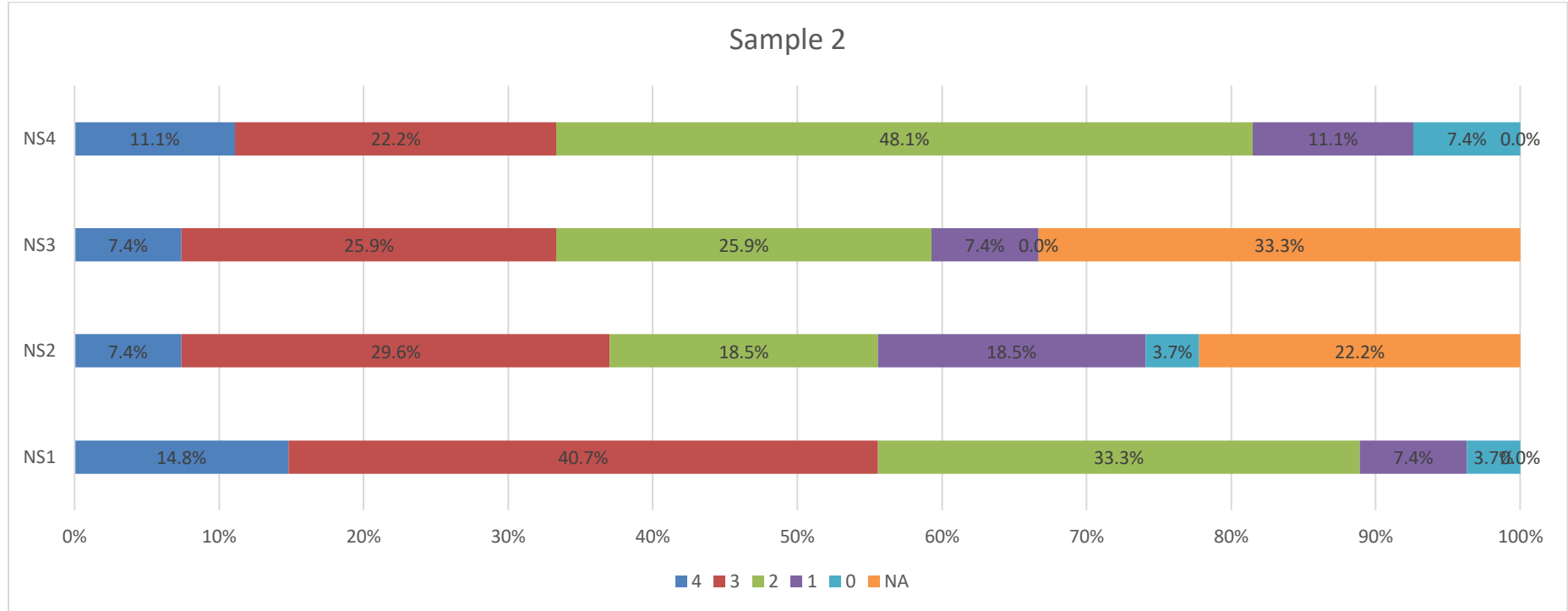
Appendix A. Results from Assessment Training for University of Louisville Natural Sciences Critical Thinking Rubric with Notes to Assessors

Sample 1



Notes: After scoring Benchmark Sample 1, assessors were given an overview of how to score the work sample based on ratings assigned by the Cardinal Core Curriculum Committee’s Assessment Subcommittee.

Sample 2



NS1 – Majority of scores were split between level 3 and level 2 performance.

(3 – “Draws conclusions that provide relevant evidence in support of a scientific model or claim. Understanding is not seriously impeded by omissions.” or 2 – “Draws conclusions, but leaves some terms undefined and/or backgrounds unknown.”)

NS2 – Unlike our first example, the student does try apply the ethogram to make a prediction about the sleeping habits of domestic cats, so it is better to score (4,3,2,1, or 0) this item instead of using the N/A rating.

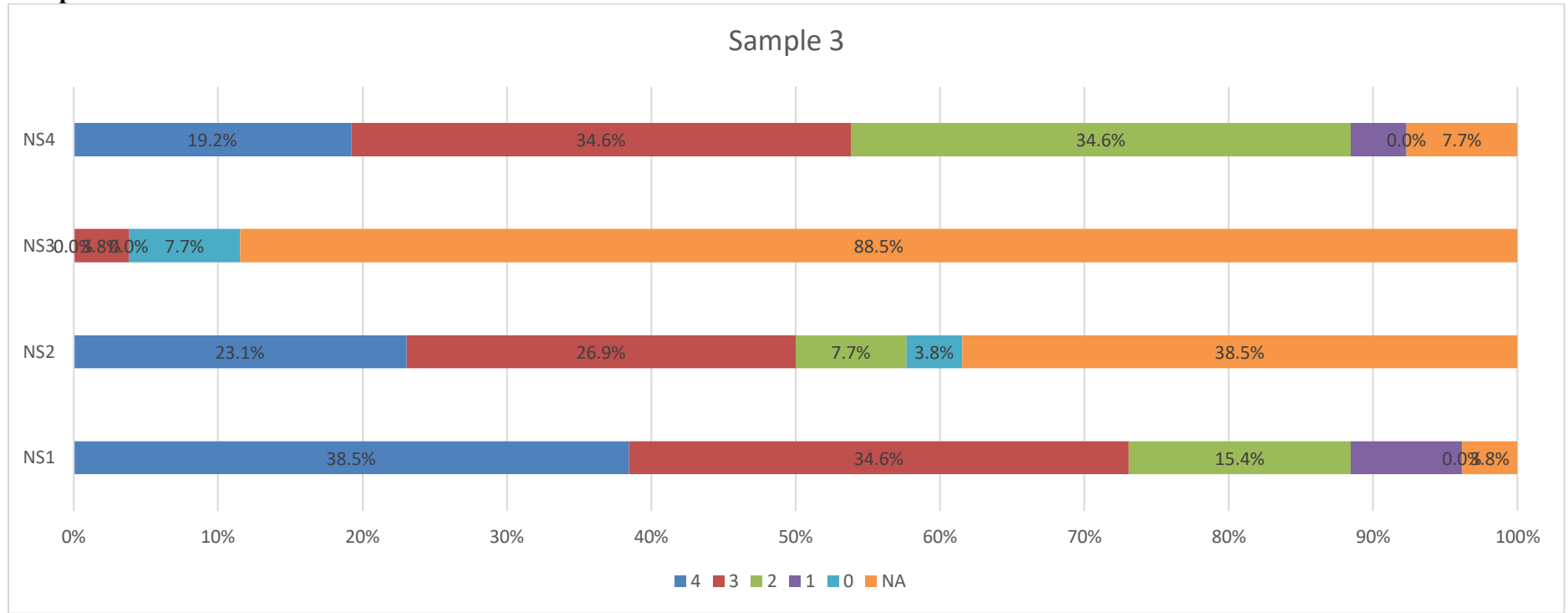
NS3 – With this item, the student does draw a conclusion about the sleeping habits of domestic cats based on the information gathered. The majority of the scores assigned (excluding N/A) were split between level 3 and level 2 performance.

(3 – “Draws conclusions which implicitly link between appropriate scientific principles and issues of personal and/or societal importance.” or 2 – “Draws conclusions which inconsistently link scientific principles to issues of personal and/or societal importance.”)

NS4 – The most common response was level 2 performance.

(2 – “Communicates basic concepts and experimental outcomes, but inconsistently uses scientific vocabulary, provides incomplete explanation of findings, and/or lacks interpretation.”)

Sample 3



NS1 – Majority of scores were split between level 4 and level 3 performance.

(4 – “Draws conclusions that provide explicit, comprehensive, and relevant evidence in support of a scientific model or claim.” or 3- “Draws conclusions that provide relevant evidence in support of a scientific model or claim. Understanding is not seriously impeded by omissions”).

NS2 – For this example, applying the appropriate calculations and providing results can be considered applying scientific principles to make predictions. Scores at level 4 or level 3 are both appropriate.

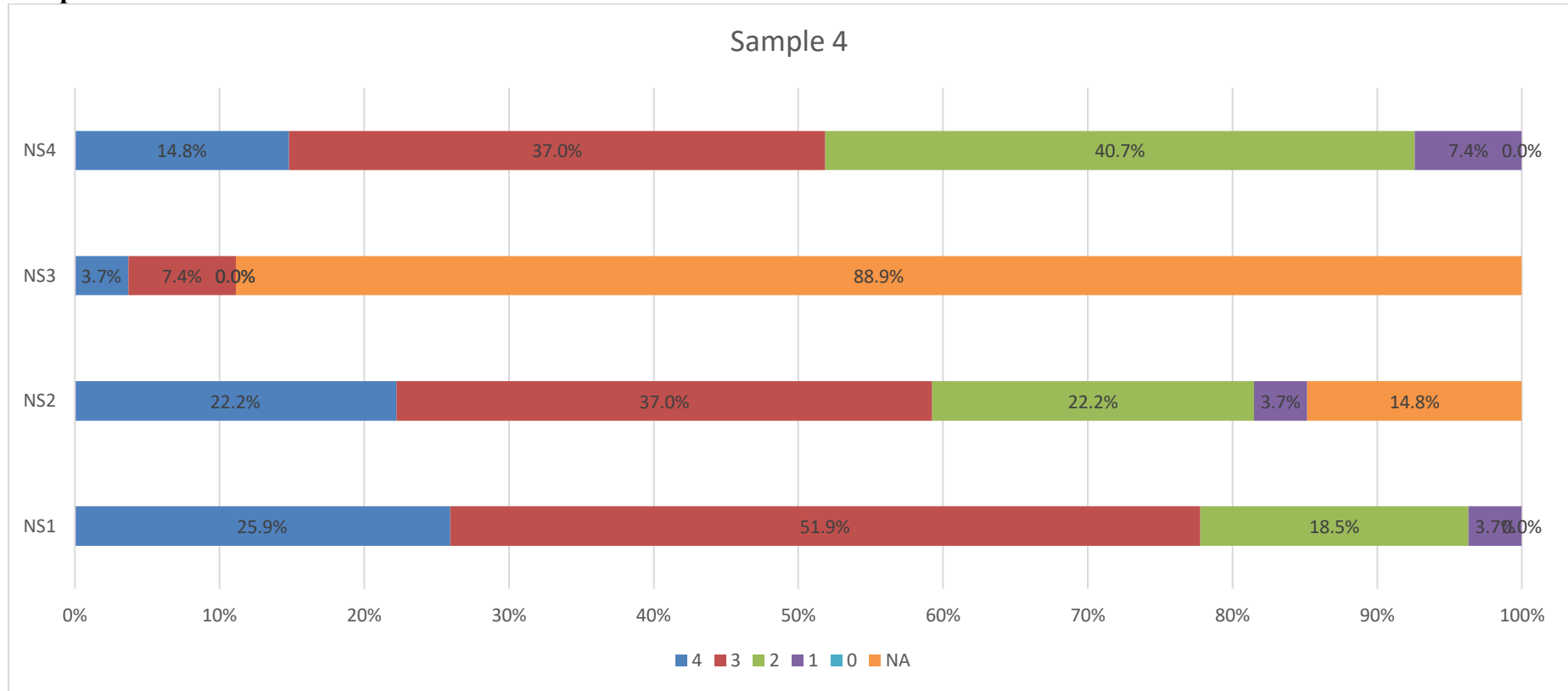
(4 – “Thoroughly applies scientific principles and sound scientific reasoning to make appropriate predictions about natural phenomena.” or 3 – “Applies scientific principles and some aspects of scientific reasoning to make predictions about a natural phenomenon.”)

NS3 – A score of N/A is appropriate for this sample since the student is not asked to connect scientific principles back to issues of personal and/or societal importance.

NS4 – Majority of scores were split between level 3 and level 2 performance.

(3 – “Communicates an understanding of scientific principles and experimental outcomes through the use of scientific vocabulary, explanations of findings, and/or interpretations, as appropriate for the discipline.” or 2 – “Communicates basic concepts and experimental outcomes, but inconsistently uses scientific vocabulary, provides incomplete explanation of findings, and/or lacks interpretation.”)

Sample 4



NS1 – Majority of scores were at a level 3 performance.

(3 – “Draws conclusions that provide relevant evidence in support of a scientific model or claim. Understanding is not seriously impeded by omissions.”)

NS2 – For this example, carrying out the experiment and applying the equations can be considered applying scientific principles to make predictions. Majority of scores assigned (excluding N/A) were at or within one level of performance level 3.

(3 – “Applies scientific principles and some aspects of scientific reasoning to make predictions about a natural phenomenon.”)

NS3 – A score of N/A is appropriate for this sample since the student is not asked to connect scientific principles back to issues of personal and/or societal importance.

NS4 – Majority of scores were split between level 3 and level 2 performance.

(3 – “Communicates an understanding of scientific principles and experimental outcomes through the use of scientific vocabulary, explanations of findings, and/or interpretations, as appropriate for the discipline” or 2 – “Communicates basic concepts and experimental outcomes, but inconsistently uses scientific vocabulary, provides incomplete explanation of findings, and/or lacks interpretation.”)