Assessment of General Physical Science Education

Phillip Vargas

Unit Level Liaison for Dept. of Physical Science

The Department of Physical Sciences is continuing to shift assessing student learning outcomes in the general education courses away from content-based assessment tools and toward process-based assessment tools. While many faculty members have adopted this philosophy years ago, the transition to having full scale adoption of a process that includes a single assessment tool to acquire this data has been long, and spring 2021 was the first semester where every general education physical science course was assessed with a process-based tool. One of the primary goals of this year's project was to develop a process that allows for this type of large-scale assessment deployment to be administered every semester.

The Assessment Tool

The most prevalent tool utilized for general education physical science courses in higher education is the Lawson Classroom Test of Scientific Reasoning (CTSR) developed by Anton Lawson in 1978 and subsequently revised until 2000. While prevalence does not necessary correspond with efficacy, it has also been extensively validated and administered to multiple institutions. This data is warehoused in the PhysPort website and is constantly being updated by participating faculty members and institutions. Administering this tool allows the department to compare our assessment results with similar courses at other universities, in addition to benchmarking them against our program-level outcomes. Capturing this data provides an excellent starting point to building a strong general education assessment foundation.

This tool measures scientific reasoning across six domains 1) conservation of matter and volume, 2) proportional thinking, 3) probabilistic thinking, 4) correlational thinking, 5) control of variables, and 6) hypothetical-deductive reasoning. These skills are essential components to science courses and are typically included when defining scientific reasoning. These sills also align with the department's program-level learning outcomes: 3) Analyze and interpret data using mathematics and computational thinking and 4) Construct explanations and engage in arguments from evidence.

Deployment System

This assessment had been administered by some professors at the section level. However, these data sets were primarily used by the instructor in appraising lesson plans and instructional techniques. In order to improve the reliability of the data and measure program outcomes, scaling the data acquisition to include more sections was necessary. This goal also aligns with the objective to move general education assessment to the departments. The challenges with this goal have been coordinating and integrating faculty members into the process. Coordinating an assessment that occurs roughly at the same time point in several sections, modalities, colleges, and teaching schedules posed a difficult undertaking. While this large-scale coordination was attempted in the past it proved cumbersome, difficult to implement and unlikely to be sustainable.

In fall 2020 the Learning About STEM Student Outcomes (LASSO) platform was piloted as a potential candidate for a large scale assessment data acquisition tool. This system was specifically designed for this type of deployment in administering, analyzing, and reporting assessment outcomes. It has 32 assessment instruments to choose from and is open source. This system was piloted with five sections across two classes.

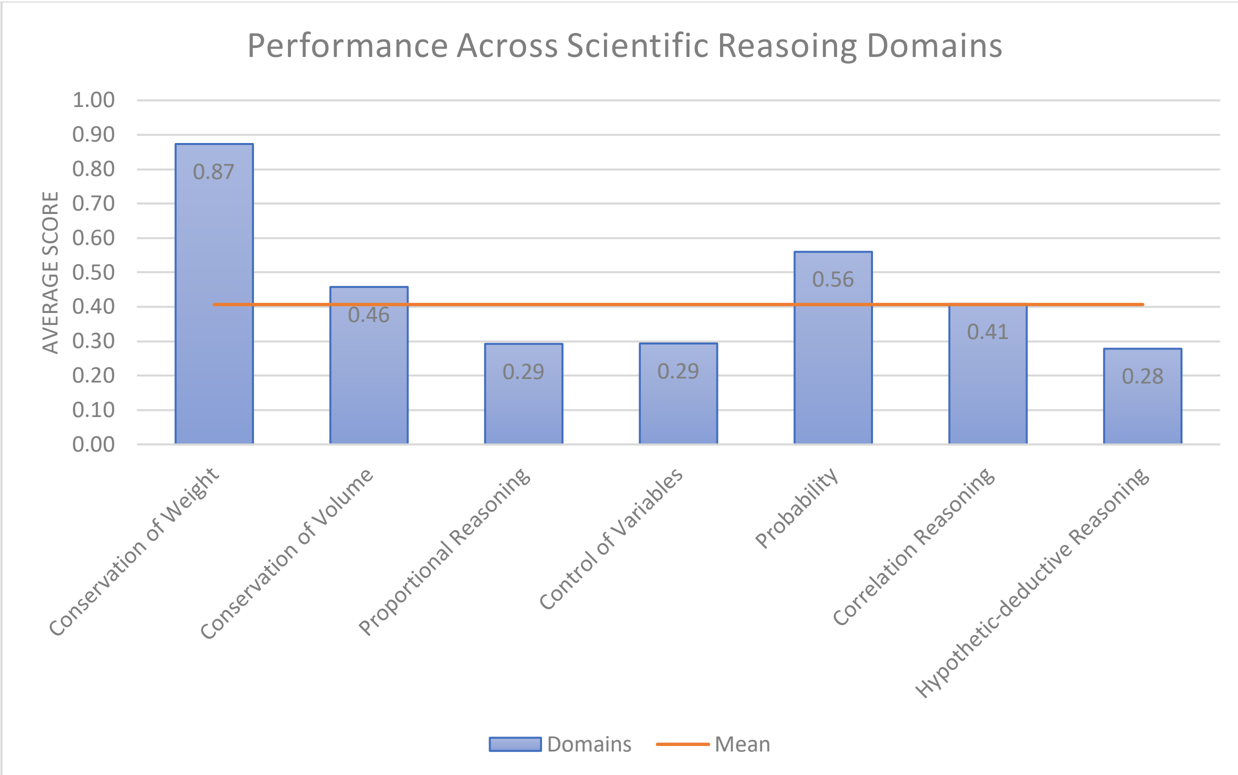
Unfortunately, the pilot was unsuccessful. At most, 1 student completed the assessment in each of the piloted sections. In informal discussions with students in the pilot sections it was reported that the emails were ignored because they were coming from a non-CCC institution, the instructions were confusing, and there were a bank of questions that did not pertain to activities they were involved in. Since this platform is part of The University of Colorado’s Learning Assistant Alliance it was developed in part to obtain comparative data between sections with and without a learning assistant, and these questions are included by default. These concerns could not be corrected within this platform, and it was ultimately decided not to be a good fit.

Although it is was clear that the LASSO system had limitations that prevented it from being incorporated into the department’s assessment process, the overarching process it used was still appealing. Therefore, a similar platform was developed within the Office 365 application suite. Utilizing Forms, Excel, Outlook, and Power Automate an instrument could be created, emailed to all of the students, and results stored in central location. Additionally, the form and email message could be customized for our students and be sent from the department liaison who could respond to questions or concerns. This system was developed and deployed in Spring 2021.

Results

In spring 2021 there were 39 sections of Physical Science classes with 850 students initially enrolled. At the time of the assessment 638 students were enrolled across these sections. A total of 82 students completed the assessment giving these results a margin of error of 10% at a confidence level 95%. While 5% margin of error is the *de facto* gold standard and the goal of the assessment committee, achieving a 10% margin of error on the first full scale deployment of a new platform was better than anticipated.

While there is not sufficient data or knowledge of confounding variables to perform comparative analysis to external institutions or internal institutional dimensions (course, modality, etc) this assessment does allow us to compare relative performance across the six scientific reasoning domains. These results illustrated some evidence that appeared contrary to anectodical evidence in the department. It is often discussed that student's scientific reasoning is higher than their mathematics background. However, in these results students initially appear to be performing relatively well in conservation and probabilistic thinking, while their understanding of how to control for variables and hypothetic-deductive reasoning is significant lower. This type of reasoning is key in experimental design and to a large part of the scientific method. While there may be significant difference between these performances in lab and non-lab classes, this point may need some reflection in the department and possibly retooling of laboratory practices. Some possible solutions could be requiring students to perform more experimental setup, develop more open-ended problems, or even potentially even having students design their own experiments.



As this instrument has been validated in several education settings, it was not necessary to analyze its efficacy. However, an analysis of independence of the performance across these domains was conducted with a correlation matrix analysis. This analysis shows any dependencies within these domains. This could illustrate that particular classroom assignments have a larger effect on a subset of these domains. This analysis showed little to no correlation across these domains and that they are independent.

	Conservation of Weight	**Conservation of Volume**	**Proportional Reasoning**	**Control of Variables**	**Probability**	**Correlation Reasoning**	**Hypothetic-deductive Reasoning**
Conservation of Weight	1.00	0.10	0.17	0.21	0.19	0.08	0.07
Conservation of Volume	0.10	1.00	0.34	0.37	0.29	-0.06	-0.02
Proportional Reasoning	0.17	0.34	1.00	0.17	0.38	0.11	0.07
Control of Variables	0.21	0.37	0.17	1.00	0.23	-0.06	0.38
Probability	0.19	0.29	0.38	0.23	1.00	0.12	0.10
Correlation Reasoning	0.08	-0.06	0.11	-0.06	0.12	1.00	-0.09
Hypothetic-deductive Reasoning	0.07	-0.02	0.07	0.38	0.10	-0.09	1.00

Additionally, this new process allows disaggregation of the data by course, session, modality, or demographics. While this type of analysis is planned for future reports, additional safeguards need to be addressed first to ensure confidentiality and anonymity.

Conclusion

In spring 2021 the first large scale administration of a single assessment instrument across all of the general education courses offered in the Department of Physical Science was conducted. This was achieved by developing an online administration and acquisition platform in the Office 365 suite. The acquired results met the 10% margin of error at a 95% confidence level and showed the strengths and weaknesses of our students learning across six domains of scientific reasoning. While students initially appear to be performing relatively well in conservation and probabilistic thinking, while their understanding in how to control for variables and hypothetic-deductive reasoning is significant lower. As this type of reasoning is key in experimental design and the scientific method, it needs to be further explored and addressed if confirmed.