

**Using Existing Assessments To Track Longitudinal Development Of  
Students' Critical And Creative Thinking Skills**

by

**Paisley Worthington**

**A Thesis  
presented to  
The University of Guelph**

**In partial fulfilment of requirements  
for the degree of  
Master of Science  
in  
Biophysics**

**Guelph, Ontario, Canada**

**© Paisley Worthington, September 2019**

## ABSTRACT

### **USING EXISTING ASSESSMENTS TO TRACK LONGITUDINAL DEVELOPMENT OF STUDENTS' CRITICAL AND CREATIVE THINKING SKILLS**

**Paisley Worthington  
University of Guelph, 2019**

**Advisor:  
Professor John Dawson**

Learning outcomes assessment (LOA) data are most commonly collected via national surveys, alumni surveys / focus groups, and locally developed surveys. While these methods collect important affective information about students and alumni, detailed data describing students' actual skills and perceived skills are needed to facilitate targeted improvements to courses and programs and enhance the LOA initiative.

I implement a triangulated LOA strategy that coordinates three essential perspectives of LOA: intention, achievement, and perception. Focussing on problem solving assessments, I develop and validate a problem solving profile, a program-level developmental rubric for problem, and a student survey. To my knowledge, the profiling tool and program-level developmental rubric for assessing problem solving are the first such tools described in the field. Using pilot data from the 2018-2019 academic year, I demonstrate how these tools and methods generate data educators were previously unable to access.

## Acknowledgements

Thank you to everyone who helped me push this research as far as I could.

Thank you John and Dale, for finding space for this project and persevering through obstacles I didn't even know were there. For encouraging me to think big, listening to my half-baked ideas, and helping me develop them into elements of my research. For giving me the freedom to take this project where I wanted it to go and for sending me all over the place to share my work.

Thank you family, for the unconditional love and support and for forcing me to take breaks when I really needed to. For being so ridiculous that I burst out laughing at random.

Thank you friends, for the good times. Cheers to beer, live music, and bike rides 🙌

## Declaration of work performed

The Problem Solving Profile Tool, the Program-Level Problem Solving Rubric, and the student survey were developed solely by Paisley Worthington. Kim Garwood, Jacqueline Hamilton, Karl Cottenie, Aron Fazekas, Joannah O’Hatnick, Tommy Mayberry, Sara Fulmer, Jeni Spencer, Claire Coulter, John Dawson, Janet Wolstenholme, and Laura Schnablegger participated in one or both of the Problem Solving Profile Tool / Program-Level Problem Solving Rubric validation meetings led by Paisley Worthington. Jeni Spencer and Aron Fazekas reviewed the anonymous student survey with Paisley Worthington before its use. Jeni Spencer and Anamika Dutta took on the responsibilities of disinterested third party for de-identification of sensitive student data. Statistical tests were selected in consultation with Lucia Costanzo. All analyses were performed by Paisley Worthington.

## Table of Contents

<b>ABSTRACT .....</b>	<b>ii</b>
<b>Acknowledgements .....</b>	<b>iii</b>
<b>Declaration of work performed .....</b>	<b>iv</b>
<b>List of Tables .....</b>	<b>viii</b>
<b>List of Figures .....</b>	<b>ix</b>
<b>Abbreviations .....</b>	<b>x</b>
<b>Chapter 1: Introduction .....</b>	<b>1</b>
<b>Keeping universities indispensable .....</b>	<b>1</b>
The call for detailed learning outcomes assessment data .....	3
<b>Learning Outcomes Assessment Perspectives .....</b>	<b>3</b>
<b>The emphasis on critical and creative thinking .....</b>	<b>6</b>
<b>Problem Solving as a measure of critical and creative thinking .....</b>	<b>8</b>
The Problem Solving Process .....	8
<b>The Current Thesis: Establishing a Triangulated, Data-driven Approach to LOA .....</b>	<b>9</b>
Research Aims.....	9
Ontology.....	11
Substantial and Original Contribution to Knowledge.....	12
<b>Chapter 2: Program Intentions (PS Scaffolding) .....</b>	<b>13</b>
<b>Supporting Students' Problem Solving Development with Authentic Assessments .....</b>	<b>13</b>
Characteristics of Problems.....	14
Inherent Problem Qualities .....	15
Problem Solving Environment .....	17
Characteristics of Problem Solvers.....	18
<b>Meaningful Assessment of Problem Solving .....</b>	<b>19</b>
Authentic and Embedded Assessments .....	21
The Problem Solving Profile Tool .....	22
<b>Methods .....</b>	<b>23</b>
Selection of Problem Solving Assignments.....	23
Development and Validation of the Problem Solving Index Tool .....	24
Using the PS Profile tool to assign AAs a PS score.....	25
Reflection.....	25
<b>Results .....</b>	<b>26</b>
PS Profiles of four Authentic Assessments .....	26
<b>Discussion .....</b>	<b>27</b>
PS Profiles of four Authentic Assessments .....	27
Advantages and Limitations.....	28
<b>Conclusion.....</b>	<b>29</b>
<b>Chapter 3: Student Achievement.....</b>	<b>31</b>

<b>Learning Outcomes Assessment Strategies</b> .....	<b>31</b>
Program Level Rubrics as an Outcomes Assessment Tool.....	34
Contextualized Outcomes Analysis .....	35
<b>Methods</b> .....	<b>37</b>
Development and Validation of the Program-Level Problem Solving Rubric .....	37
Alignment of Program-level PS rubric with Institutional Critical and Creative Thinking Rubric .....	38
Alignment of AA rubric with Program-level PS rubric.....	38
Collection of Student Performance Data.....	38
Analysis of Student Performance Data .....	39
Reflection.....	40
<b>Results</b> .....	<b>40</b>
Rubric Structure.....	40
Participation.....	41
Alignment of the PS Rubric with AA rubrics .....	42
Student Performance .....	43
<b>Discussion</b> .....	<b>44</b>
Alignment of the PS rubric with AA rubrics .....	44
Student Performance .....	46
Advantages and Limitations.....	48
<b>Conclusion</b> .....	<b>49</b>
<b>Chapter 4: Student Perceptions</b> .....	<b>51</b>
<b>Significance of the Student Perspective</b> .....	<b>51</b>
Surveys as Diagnostic Tools .....	51
<b>Methods</b> .....	<b>52</b>
Development and Validation of the Student Survey.....	53
Administration of the Survey.....	54
Analysis Techniques .....	55
Reflection.....	56
<b>Results</b> .....	<b>57</b>
Participation.....	57
Validity.....	57
Association of PS with critical and creative thinking .....	58
Perceived ability to communicate PS and critical/creative thinking skills .....	58
Perceived effect of selected AAs .....	60
Student perceptions over time.....	62
<b>Discussion</b> .....	<b>64</b>
Validity and Association of PS with critical and creative thinking.....	64
Perceived ability to communicate PS and critical/creative thinking skills .....	65
Perceived effect of selected AAs .....	65
Student Perceptions over time .....	67
<b>Conclusion</b> .....	<b>69</b>
<b>Chapter 5: A More Holistic View of an Undergraduate Science Program</b> .....	<b>70</b>
<b>Considering the Triangulated Data</b> .....	<b>70</b>
<b>Future Work</b> .....	<b>72</b>

<b>Demonstrating indispensability using data.....</b>	<b>73</b>
<b><i>References</i> .....</b>	<b>74</b>
<b><i>Appendices</i> .....</b>	<b>81</b>
<b>Appendix A: The Problem Solving Profile Tool .....</b>	<b>81</b>
<b>Appendix B: The Program-Level PS Rubric .....</b>	<b>82</b>
<b>Appendix C: Student Perceptions Survey .....</b>	<b>84</b>

## List of Tables

Table 1: Three categories of characteristics that contribute to the difficulty of a problem. .....	15
Table 2: Traits of structured and unstructured problems. ....	16
Table 3: Overall PS Profile scores of four authentic assessments employed in the Microbiology degree program at the University of Guelph.....	26
Table 4: Number of students who consented to share their grades on the selected AA. .....	42
Table 5: Student participation in the anonymous survey .....	55
Table 6: Spearman's Rank Order Correlation for paired survey questions. ....	58
Table 7: Kruskal-Wallis H-test values for survey question responses across year levels. .....	64
Table 8: Paired survey questions used to gauge validity of student responses. ....	85
Table 9: Equivalent survey questions used to compare student perceptions of PS and critical and creative thinking. ....	85

## List of Figures

Figure 1: Projected generational distribution of Canada's working population in 2031....	2
Figure 2: The Problem Solving Profiles of four authentic assessments employed in the Microbiology degree program at the University of Guelph.....	26
Figure 3: Linear alignment approach to learning outcomes assessment.....	31
Figure 4: Integrated alignment approach to learning outcomes assessment. ....	33
Figure 5: Program-level PS Rubric quadrants targeted by selected assignments of the University of Guelph MICR program.....	43
Figure 6: Student performance on PS aspects of selected AAs according to a) PS Profile Score and b) year level. ....	44
Figure 7: Cross-tabulations of student responses on paired survey questions probing perceived ability to communicate a) problem solving skills and b) critical and creative thinking skills. ....	60
Figure 8: Cross-tabulations of student responses on paired survey questions probing perceived effect of selected AA on a) problem solving skills and b) critical and creative thinking skills. ....	62
Figure 9: The anonymous student survey administered in this study. ....	84

## Abbreviations

AA = authentic assessment

LO = learning outcome

LOA = learning outcomes assessment

PS = problem solving

PSP = Problem Solving Profile

MICR = in reference to the University of Guelph undergraduate Microbiology Honour's degree

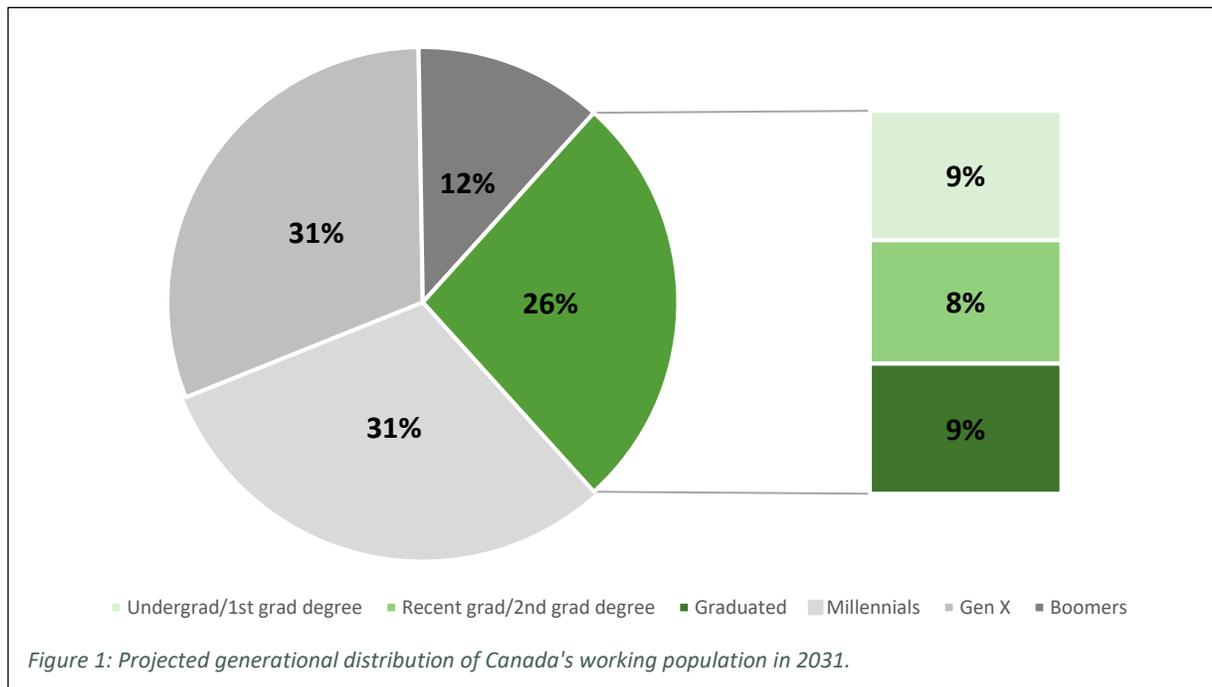
## Chapter 1: Introduction

### Keeping universities indispensable

Post-secondary education has, for decades, been viewed unquestionably as an investment for students wishing to develop sophisticated cognitive skills and to establish a well-respected vocation. However, in the last few years, technology has opened the door for companies to redefine the knowledge sector by providing low-cost, accessible learning opportunities (ex., Khan Academy, Coursera, edX, YouTube, among others) (“How has technology changed education?,” 2019). Skeptics of higher education use the Internet to publish their doubts about the true value of a university education (ex., Blackwell, 2018; Hannah, 2019; Ruppel Shell, 2018; Sodha, 2018). Despite popular articles explaining the various benefits of pursuing higher education (ex., “Higher education still worth the money, new research suggests,” 2016; “Is university worth the money?,” 2018; Newton, 2018), the presence of two starkly different opinions in the media can cause potential university students to reconsider their educational pathway. For post-secondary education to remain relevant, ambassadors of higher education must communicate how university education equips students with essential skills; for instance, teaching cognitive skills like problem solving.

The generation following the Millennials will need highly proficient cognitive skills, perhaps more than any other generation. Globally, we have massive problems that are the result of short-sighted thinking: the energy crisis, air and water pollution, deforestation, the Great Pacific garbage patch, and ethical dilemmas surrounding day-to-day human rights and medical advancements. Some scientists predict that the environmental problems must be addressed by the year 2030 or the environment will be

irreversibly affected by runaway climate change (Steffen et al., 2018; Walsh et al., 2017; World Economic Forum, 2017). Based off the 2016 Canadian census data, approximately 26% of the working population in 2031 will be younger than the Millennials (Figure 1) (Statistics Canada, 2017).



These individuals will have to confront the world's issues and will require well-refined critical and creative thinking skills – skills that universities are increasingly instilling in their students via degree-level expectations (ex., Desmarais, 2012). In 2019, the ages of these individuals range from 8 – 22 (Statistics Canada, 2017). To encourage the next generation to enrol in post-secondary education and acquire advanced critical and creative thinking skills, institutions must show their relevance and indispensability. One strategy to articulate an institution's educational value is to leverage learning outcomes assessment data (Ewell, 2009).

## The call for detailed learning outcomes assessment data

The learning outcomes assessment (LOA) movement sweeping the post-secondary sector both nationally and globally demands that colleges and universities use evidence to substantiate their students' learning gains (Goff et al., 2015). As of 2015, 30% of Canadian universities indicate that formal learning outcomes are defined in all of the institution's departments (MacFarlane & Brumwell, 2016). Behind program accreditations, the second most motivating factor for Canadian universities to engage in LOA is the institutions' commitment to continuously improve their courses and programs (MacFarlane & Brumwell, 2016). LOA data are most commonly collected via national student surveys, alumni surveys / focus groups, and locally developed surveys (MacFarlane & Brumwell, 2016; Williams, 2014). While these methods collect important affective information about students and alumni, they are not well-suited to accurately measure students' actual achievement and they are often disconnected from the classroom where student skills are being developed (Campbell & Cabrera, 2011; Gordon, Ludlum, & Hoey, 2008; Porter, 2011, 2013). Detailed LOA data describing students' performance on specific tasks throughout the program are needed to facilitate targeted improvements to courses and programs and enhance the LOA initiative (Goff et al., 2015; Klemenčič & Chirikov, 2015).

## Learning Outcomes Assessment Perspectives

Stewards of higher education engage in LOA most often to serve one of two purposes: to provide data for accountability purposes or to improve programs (Ewell, 2009). The motivation behind a LOA initiative impacts the methodology of the project.

LOA for accountability (such as accreditation) often adopts standardized methods to produce standardized data that are prescribed by the reviewing authority (such as an accreditation board) (Ewell, 2009). In these scenarios, the measures of post-secondary programs are often outlined explicitly by the reviewing authority. LOA for improvement, however, can take advantage of a triangulated methodology, collecting varied information from multiple tools and processes (Ewell, 2009). Educators using LOA for improvement have the benefit of selecting institutional outcomes of interest to study in depth, defining in-house standards of success, and drawing from real assessments used in the program as data sources (Ewell, 2009). Given the flexibility and freedom associated with LOA for improvement, researchers should consider the degree of variation and redundancy required for a high quality LOA methodology that is still feasible. My thesis examines LOA for improvement from three specific perspectives: 1) the program's intentions (Chapter 2), 2) student achievement of LOs (Chapter 3), and finally 3) student perceptions of their education (Chapter 4).

Researchers conducting a LOA for improvement may consider the program's intentions for student learning. Ewell argues that in any LOA initiative, assessments placed at notable positions within an undergraduate program ought to be emphasized as indicators of longitudinal student development (Ewell, 2009). Instructors of King's College scaffold capstone assessments throughout the degree such that the assessments become more challenging with each year level – this scaffolding represents the deliberate calibration of the faculty's expectations of their students (Ewell, 2009). Thus, a shared understanding of the program's intentions is a LOA perspective that can be used to contextualize student performance data.

Student achievement data provide another important perspective in LOA initiatives. Most LOA projects look to student achievement data to determine whether learning outcomes are being met by students. LOA for accountability purposes often make use of standardized assessments that are developed externally from the institution (ex. MCAT or LSAT testing); tests such as these are used to determine whether a student has objectively gained particular knowledge or skills. While these tests serve their purpose, they are ill-suited for LOA for improvement projects, where detailed and personalized data about a specific program at a specific school can provide useful data that is interpreted and acted upon more readily. As suggested by Ewell (2009), assessments employed in the program of interest may be used to represent the student achievement perspective in a LOA project.

A third perspective to include in a LOA for improvement project is the students' perception of their education. Increasingly, value is placed on the student voice as a source of feedback for teaching and learning processes (ex. Beattie, 2012; Cook-Sather, Bovill, & Felten, 2014; Delpish et al., 2010; Fielding, 2001; McAlpine & Asgharb, 2010; O'Neill & McMahon, 2012; Thiessen, 2006). One study recounts how researchers engaged undergraduate students in the curricular overhaul of a physiotherapy program (O'Neill & McMahon, 2012) while another group describes their experience working with doctoral students to co-construct departmental practices and policies concerning the doctoral program (McAlpine & Asgharb, 2010). Students are the primary recipients of undergraduate programs and have the potential to provide an insightful recollection of how the program is actually experienced. Thus, some representation of the student voice ought to be included as an additional perspective in a thorough LOA project.

## The emphasis on critical and creative thinking

Critical and creative thinking skills are crucial to thrive in our constantly changing global community and are a common target of LOA initiatives (“Employability Skills,” 2018; Gini-Newman & Case, 2015; Puccio, Mance, & Murdock, 2011; Simper, Frank, Scott, & Kaupp, 2018). In the Herbert report (2015), the Ontario government listed critical thinking as one of three learning outcomes that should be emphasized in higher education programs in Ontario (Herbert, 2015). However, the report also states that “these [skills] are not transparently or consistently measured, assessed, or validated across the system” (p. 38), suggesting that further research concerning critical and creative thinking is required (Herbert, 2015). Presumably student thinking is improving, but until recently neither longitudinal performance data nor a framework for understanding critical and creative thinking development existed prior to 2018 as evidence to support this claim, when one study from Queen’s University was released (Simper et al., 2018).

Researchers from Queen’s University launched a four-year longitudinal study examining the development of critical thinking, problem solving, communication, and lifelong learning in numerous undergraduate programs representing engineering, science, social science, and humanities (Simper et al., 2018). The study used multiple data sources, including standardized tests, assignments graded with VALUE rubrics, and interviews with students and instructors to collect learning progression data about its students. Researchers found heavy overlap with the component skills of the cognitive skills being measured, namely critical thinking, problem solving, and communication (Simper et al., 2018). While this blending of skills did not hinder the project’s success in

collecting summative information concerning students' academic achievement, a clear distinction of skills is required if institutions wish to use this type of data formatively to examine curriculum mapping and identify areas for improvement.

With this reasoning, my study focusses solely on critical and creative thinking. Although critical and creative thinking are unquestionably important outcomes of post-secondary education, the higher education community lacks a definition of this skill that is universally accepted by its practitioners (Deller, Brumwell, & Macfarlane, 2015). I adopted a different conceptualization of complex cognitive skills than the group from Queen's. Instead of treating all cognitive skills equally and independently, in my model these skills are in a hierarchy, with critical and creative thinking as intertwined skills and the foundation from which other skills stem (including problem solving, communication, literacy, ethical behaviour, etc.) (Gini-Newman & Case, 2015; Mulnix, 2012; Paul, 1993).

Since critical and creative thinking is treated as a foundational skill, discussing critical and creative thinking independently of context is difficult (Deller et al., 2015; Mulnix, 2012). Further, since multiple theories present competing definitions of critical and creative thinking, explicitly teaching this skill in their courses is challenging for instructors ("Critical Thinking," 2018; Martini & Clare, 2014; Mulnix, 2012). As a result, students may not be aware of their transferrable critical and creative thinking skills or may struggle to communicate their abilities to employers (Edge, Martin, & Mckean, 2018; Herbert, 2015; Martini & Clare, 2014). In addition, assessing critical and creative thinking is difficult primarily because we cannot assess thought directly; instead, we must assess the quality of a product and make inferences about the depth of thinking

used to create the product (Holmes, Wieman, & Bonn, 2015; Jonassen, 2000; Maki, 2010; Mayr, Smuc, & Risku, 2010).

One way to respond to the challenges of assessing critical and creative thinking is to consider these skills in the context of activities that rely heavily on them, for instance, problem solving (Mayr et al., 2010). At the University of Guelph, Problem Solving (PS) is one component of the Critical and Creative Thinking institutional learning outcome (Desmarais, 2012). Using PS as an indicator of critical and creative thinking is logistically advantageous because students can produce concrete evidence that demonstrate their ability to solve a problem. Situating PS as an indicator of critical and creative thinking therefore allows us to indirectly gauge a student's ability to think (Frederiksen & Ward, 1978; Mayr et al., 2010; Pithers & Soden, 2010). Given the relationship between critical and creative thinking and PS, I will next examine the literature surrounding PS and relate that knowledge to practices held in a Canadian undergraduate biology degree.

## Problem Solving as a measure of critical and creative thinking

### The Problem Solving Process

Problem solving is a higher order cognitive process that integrates disciplinary knowledge with metacognitive skills. To successfully solve a problem, a thinker must have both strong disciplinary knowledge to understand the context of the problem and sophisticated metacognitive skills to regulate their thinking as they navigate the different stages of solving a problem (Greiff, Fischer, Stadler, & Wüstenberg, 2015; Jonassen, 2000; G. Klein et al., 2003; Mayr et al., 2010).

Numerous PS theories offer conceptualizations of the PS process that bear resemblance to one another and can often be broken into four main stages: 1) understanding the problem, 2) exploring potential solutions, 3) implementing a selected solution, and 4) reflecting on the experience (ex. Bransford & Stein, 1984; Bryant, 2006; Ilevbare, Probert, & Phaal, 2013; Matsuo & Nakahara, 2013; McLennan & Omodei, 1996). General PS models have been criticized for being overly reductive, sacrificing situational nuances and the fluid nature of complex problems for the sake of producing a generalized PS model (Jonassen, 2000; O'Loughlin & McFadzean, 1999). General PS models provide limited insight into the PS process because, in real life, problem solvers must adapt their PS strategy to each problem and are often required to consider disciplinary nuances to find a successful solution. Although solutions are often rooted in disciplinary context, a student's PS ability can also be evaluated from a metacognitive perspective independent of discipline. Viewing PS activities through a metacognitive lens allows educators to draw attention to and emphasize the transferrable aspects of PS that might be overlooked in favour of disciplinary knowledge (Edge et al., 2018; G. Klein et al., 2003; Martini & Clare, 2014).

### The Current Thesis: Establishing a Triangulated, Data-driven Approach to LOA Research Aims

It is clear that we require data to inform learning outcomes assessment discussions. The goal of my research was to develop a sustainable method of collecting such data; specifically, about how students develop critical and creative thinking skills through PS assignments. I accomplished this goal by examining three essential perspectives of

learning outcomes assessment: intention, achievement, and perception. My method addresses these research questions:

**Research Question 1:** Is an undergraduate science program employing more advanced PS assignments in senior years? (intention)

**Research Question 2:** Are students' critical and creative thinking abilities progressing through an undergraduate science curriculum? (achievement)

**Research Question 3:** How do undergraduate science students perceive their developing transferrable critical and creative thinking skills? (perception)

To meet my research goal, I establish a learning outcomes assessment strategy referred to as "The Framework." The Framework uses 1) a Problem Solving Profile (PSP) tool to characterize the challenge of PS assignments (intention), 2) a program-level developmental rubric outlining the PS process and defining success criteria thereof (achievement), and 3) a student survey to collect students' perceptions of their own PS development (perception).

All three tools (PSP tool, developmental rubric, and student survey) are researcher-generated. To the best of my knowledge, The Framework is original and has not been established elsewhere. The principles of The Framework are transferrable to any undergraduate context.

In the 2018-2019 academic year, I implemented a pilot of The Framework and collected data concerning 1) assignments, 2) student performance, and 3) student perceptions. Data regarding one assignment from each year of the Microbiology degree program were collected and analyzed. This information was used to describe the

relationships between the program, assignments, and students and answer my research questions.

## Ontology

I approached this study from a critical realist and contextual standpoint, using multiple data sources to approximate how critical and creative thinking develops through an undergraduate science degree. Critical realism is an appropriate ontological framework for this study because critical and creative thinking development is complex and cannot be assessed fully from one perspective; multiple and varied sources of information are needed to obtain an accurate understanding of how students are learning to think critically and creatively. Research implemented through a contextual lens considers the context within which information is produced to appropriately interpret the data (Braun & Clarke, 2013). Combining critical realism and contextualism in my approach provides space to embrace the transferable nature of critical and creative thinking; representations of student thinking will vary depending on the assignment being examined. Further, an authentic evaluation of student learning considers multiple dimensions of learning outcomes assessment; in this study: intention, achievement, and perception. Incorporating three different methods of collecting empirical and affective data allows for a data-driven, outcomes-based analysis of how students' PS skills are supported and developed by the program, and how the skills are understood by the students themselves. Principles of learner-centeredness guides the positioning of this study; this work is oriented to generate real data that considers the learning environment provided to students, students' success in those learning

environments, and students' understanding of how their experience did or did not have a positive impact on their growth.

### Substantial and Original Contribution to Knowledge

This work will be foundational for future studies examining longitudinal critical and creative thinking achievement in students. This study saw the development and validation of two tools: a tool used to characterize the PS profile of undergraduate assignments and a program-level developmental rubric outlining the stages of the PS process and describing different levels of achievement thereof. To my knowledge, these are the first such tools produced and piloted to analyze and assess PS through a program. This work produced information that can guide college-wide discussions about curriculum development, program mapping, and students' perceived value of post-secondary education (Goff et al., 2015; Pollock et al., 2018).

## Chapter 2: Program Intentions (PS Scaffolding)

### Supporting Students' Problem Solving Development with Authentic Assessments

Education often emphasizes isolated, simple, well-structured problems even though students will also encounter more challenging, contextual problems in their careers (Jonassen, 2000; Newmann, Lopez, & Bryk, 1998). Students who are able to solve artificial problems in school are not necessarily equipped to solve 'messy' problems that they will surely encounter after graduation because these two types of problems demand different thinking skills (Jonassen, 2000; G. Klein et al., 2003; Newmann et al., 1998). Problems that evoke effective learning require the use of metacognitive skills because they are usually ill-structured, complex, and are situated in an authentic setting (G. Klein et al., 2003; Quintana et al., 2009). Further, by engaging students in PS activities educators can simultaneously support students' PS development and obtain the most up-to-date status of students' PS skills (Joordens, Paré, & Collimore, 2014). Given these revelations, a student's PS abilities cannot be accurately appraised by a standardized test, but rather by activities that demand students to exercise authentic PS skills (Joordens et al., 2014; Maki, 2010; Newmann et al., 1998).

The learning outcomes assessment movement encompasses more than simply measuring student achievement of outcomes. For a more thorough understanding of the longitudinal development of students' PS skills, we must consider how programs scaffold and position assignments to promote sustained learning. The extent of a learner's PS development is determined by what is required for the problem at hand (Mitchell et al., 2005; Wood, Bruner, & Ross, 1976). To promote the development of increasingly sophisticated PS skills, assignments should pose more advanced problems

as students progress through their degrees. Therefore, a method to discern between different challenge levels of PS assignments is required to allow instructors to consider how PS challenges are scaffolded over the course of an entire program.

Using literature reviewed below, I have developed a tool that characterizes a problem solving profile (PSP) for individual assignments. This tool encourages instructors to document the inherent challenge of the PS assignments they administer to students while considering the context within which the assignment is delivered.

### Characteristics of Problems

A problem can be defined as a situation where the action sequence required to change the current state of a scenario to some desired end state is unknown (Jonassen, 2000). This conceptualization encompasses diverse problems ranging from formulaic mathematical exercises to larger social justice issues. Problems have previously been categorized according to dichotomies, such as routine/non-routine, structured/unstructured, closed-ended/open-ended (ex. Puccio et al., 2011), but such basic classifications face criticism for oversimplifying problems by overlooking the many complexities and contextual nuances that contribute to the problem (Jonassen, 2000).

Numerous factors contribute to how well a problem can be solved by a particular student. These factors can be grouped into one of three categories: inherent qualities of the problem, PS environment, and individual differences in problem solvers (Table 1) (Jonassen, 2000). Through assignments, educators can control the inherent qualities of a problem and the PS environment to stimulate PS development in our students.

*Table 1: Three categories of characteristics that contribute to the difficulty of a problem (from Jonassen, 2000).*

<b>Inherent Problem Qualities</b>	<b>Problem Solving Environment</b>	<b>Characteristics of Problem Solvers</b>
<ul style="list-style-type: none"> <li>- Structure</li> <li>- Complexity</li> <li>- Abstractness</li> </ul>	<ul style="list-style-type: none"> <li>- Clear Expectations</li> <li>- Cues/Feedback</li> </ul>	<ul style="list-style-type: none"> <li>- Disciplinary Knowledge/Reasoning</li> <li>- Structural/Procedural Knowledge</li> <li>- General PS Strategies</li> <li>- Confidence</li> <li>- Motivation/Perseverance</li> </ul>

### Inherent Problem Qualities

#### *Structure*

The structured/unstructured paradigm of problems considers how varying degrees of uncertainty and correctness influence the PS process, with structured problems possessing a high level of certainty and correctness and unstructured problems lacking such a design (Table 2) (Jonassen, 2000). Although real life poses both structured and unstructured problems, the latter is underemphasized in higher education (Jonassen, 2000). The prominence of structured problems in education may persist based on the belief that the processes of solving structured and unstructured problems are the same, an assumption that has been debunked by studies showing that solving structured problems requires different skills than solving unstructured problems (Brookhart, 2010; G. Klein et al., 2003; Newmann et al., 1998).

*Table 2: Traits of structured and unstructured problems.*

<b>Structured</b>	<b>Unstructured</b>
<ul style="list-style-type: none"> <li>- all problem elements are presented</li> <li>- prescriptive solution</li> <li>- involves predictable application of a finite number of well-known rules or techniques</li> <li>- solutions are knowable (little uncertainty)</li> </ul>	<ul style="list-style-type: none"> <li>- incorporates uncertainty</li> <li>- may possess multiple solutions or no solutions</li> <li>- solution requires learner to use discretion in deciding which knowledge is relevant to the problem</li> <li>- may require learners to form and express personal judgments</li> </ul>

### *Complexity*

A problem's complexity is related to the number of variables and their dynamic, interconnected relationships within the problem (Jonassen, 2000). Problem difficulty increases with problem complexity because complicated problems present a larger cognitive load to learners (van Merriënboer, Kirschner, & Kester, 2003). Typically unstructured problems tend to be more complex than structured problems, but there are many examples of both simple unstructured problems (for example purchasing a vehicle) and complex structured problems (for example role play video games) (Jonassen, 2000).

### *Abstractness*

A problem's abstractness is related to the degree to which the problem is rooted in some context (Jonassen, 2000). Students with no prior background knowledge are more apt to successfully solve contextualized problems than abstract problems; however, students do perform better when they have prior disciplinary knowledge (Mulder, Lazonder, & de Jong, 2010). Moreover, recent literature considers PS skills

discipline-specific because problems “are situated, embedded, and therefore dependent on the nature of the context or domain” and solving such problems enables students to develop discipline-specific reasoning skills (p. 68, Jonassen, 2000). While this point stands, proficient problem solvers working through a situated problem still rely on non-disciplinary metacognitive skills to regulate their thinking (G. Klein et al., 2003; Mulder et al., 2010; Quintana et al., 2009; Reiser, 2004). This notion is demonstrated by the fact that skilled problem solvers can successfully solve problems in a new discipline, provided they acquire the appropriate domain knowledge (Mulnix, 2012). The problem solver is able to examine the underlying structure of the problem at hand and is equipped to “identify and formulate questions relevant” to the problem (Mulnix, 2012, p. 470).

## Problem Solving Environment

### *Clear Expectations*

To successfully complete an assignment, students must understand what constitutes good performance on said assignment (Biggs, 1996). Assignments that do not set out a clear goal for students often yield vague responses that are off-target (Simon, 1988). Simon suggests instructors design assignments to be clear, complete, and candid, and points out that in doing so, instructors are not ‘doing’ the assignment for the students but, rather, adopting a best practice (Simon, 1988). As students become more trained in their discipline, the necessary length and depth of instructions needed to explicitly communicate assignment expectations may change. For example, a first-year student with no experience finding reputable sources or writing proper

literature reviews will need more direction on a literature review assignment than a fourth year student performing a similar task. However, to maximize success, both students must have a clear understanding of what is expected of them.

### *Cues/Feedback*

Instructors can embed cues or hints into their assessments to nudge students towards a successful solution (Jonassen, 2000). Feedback practices are important in communicating to students how they can improve their learning and successfully meet course expectations (Banta, Lund, Black, & Oblander, 1996). Instructors can use cues and feedback in assignments to scaffold assignments and help novice problem solvers navigate ambiguous problems (Schmidt, Rotgans, & Yew, 2011). The principle behind scaffolded education is to support students improving their PS skills until they are able to complete challenging assignments with little to no assistance, thereby meeting program expectations (Goff et al., 2015; van Merriënboer et al., 2003). The level of support that should be built into an assignment depends on a number of factors including student competence, course context, and the inherent challenge of the problem.

### *Characteristics of Problem Solvers*

Beyond the inherent qualities that contribute to a problem's difficulty, learners themselves also bring variance in their ability to solve problems (Mulder et al., 2010). Previous experience with solving problems drastically affects a person's ability to solve new problems; experienced problem solvers recognize situational circumstances and

use their memory to match the current problem to one they have successfully solved in the past (Ross, Klein, Thunholm, Schmitt, & Baxter, 2004). While expert problem solvers possess dense, interconnected knowledge organizations and tend to find adequate solutions quickly (Ambrose, Bridges, DiPietro, Lovett, & Norman, 2010; Mulder et al., 2010; Ross et al., 2004), some literature questions whether the solution they select is the best solution to implement (Gary Klein, Associates, & Ara, 2008). Novice problem solvers tend to have sparse, disconnected knowledge organizations, so they solve problems by progressing slowly and deliberately through each of the PS stages, forging connections as they proceed (Ambrose et al., 2010; Mulder et al., 2010). Of the three categories that contribute to the difficulty of a problem (Table 1), *Characteristics of Problem Solvers* is the only category that is out of educators' – and researchers' – control. Although the population of students at the University of Guelph bring diverse experiences to the classroom, I assume that the undergraduate students participating in my study are novice problem solvers who are well positioned to expand their PS skills by studying a discipline and developing their metacognitive abilities.

### Meaningful Assessment of Problem Solving

Although there are currently no tools to formally rank PS assignments, academics have studied the impact of student assignments on PS ability. Some studies have investigated the impact of high-quality assignments on student performance in Chicago elementary schools (ex. Bryk, Nagaoka, & Newmann, 2000; Mitchell et al., 2005; Newmann et al., 2001, 1998). In these projects, researchers collected a sample of assignments and used researcher-generated rubrics to evaluate the intellectual

demands posed to students by assessing the assignment's ability to engage students in knowledge construction, communication, and real life application of principles (Mitchell et al., 2005; Newmann et al., 1998). Although one study presented their criteria within disciplinary contexts, Newmann et al. deliberately kept their criteria content-neutral to focus on students' transferrable skills and ensure that their standards could be used to evaluate the intellectual quality of assessments across disciplines (Newmann et al., 1998).

The concepts of contextualized analysis of student performance and discipline-neutral evaluation of assignments are pertinent when considering how learning outcomes – including PS – can be meaningfully assessed in an undergraduate degree program (Klemenčič & Chirikov, 2015). While the work done in Chicago elementary schools highlights the importance of interpreting student performance data within context, the rubrics assessing assignment quality were subject to interpretation by the randomly selected evaluators participating in the study (Bryk et al., 2000; Newmann et al., 1998). In these studies, researchers examined standardized elementary school curricula, permitting the reasonable use of randomly selected assessment evaluators (Bryk et al., 2000; Mitchell et al., 2005; Newmann et al., 1998). With the exception of accredited programs, post-secondary curricula are not standardized, thus any efforts to implement an evaluation of assignment quality using blind, randomized reviewers would be misguided.

## Authentic and Embedded Assessments

Given that PS is most effectively evaluated in discipline-specific contexts where students are pushed to their intellectual limits, authentic assessments (AAs) are ideal tools to collect information about students' PS abilities because these assessments are complex, meaningful tasks that require students to use their knowledge and skills in an original way (Newmann et al., 2001). AAs are based on real life activities, align with high-level learning outcomes (ex. program-level and institutional-level), and encourage students to focus on long-term growth (AACU, 2017; Newmann et al., 1998; Newmaster, Lacroix, & Roosenboom, 2006; "Signature Assignments and Rubrics," 2016; Svinicki, 2004; Wiggins, 1990). Because AAs can be linked to a cascade of learning outcomes, student performance data from assessments embedded in a degree program can be used to make inferences about how students are achieving high-level learning outcomes (Maki, 2010; Martini & Clare, 2014; Worthington, Dewancker, LaRush, Lackeyram, & Dawson, 2017). The term "embedded" is used to describe assessments that are built into a course (Wilson & Sloane, 2010), while the term "authentic" refers to the context of the assessment and how closely the assessment context aligns with both the real-world context and the student's capabilities at the time of assessment (Newmann et al., 1998; Newmaster et al., 2006; Svinicki, 2004; Wiggins, 1990). Assessments can be embedded and authentic simultaneously. I selected four embedded AAs (one from each year of a four-year degree program) for this study in an attempt to capture the students' most sophisticated PS skills at various stages of learning in the program.

## The Problem Solving Profile Tool

The recognition that PS is a measurable indicator of critical and creative thinking makes assessment of student thinking possible. PS is a cognitive process that both transcends discipline and is rooted in discipline; this complexity points to well-designed assessments as tools that can simultaneously support and evaluate the metacognitive aspects of students' PS skills. However, there are currently no published tools that compare the PS content of authentic assessments. To the best of my knowledge, this thesis offers the first published tool designed to characterize the PS challenge of various assignments. The development of this tool was guided by the literature, which highlights how problem attributes, assignment design, and student experience influence PS assessments.

The nature of an assignment can limit the depth of thinking a student engages in when completing the assignment and, by extension, our ability to gauge a students' true cognitive abilities. For example, a primary school-level physics test would not be an appropriate tool to gauge the depth of Dr. Stephen Hawking's true genius. To the best of my knowledge, I pioneered the development of the first PS classification tool to evaluate the PS demands posed by assignments used in an undergraduate program. The classification tool follows the format of an index, which is an appropriate representation in this project because indices relate numerous qualities to produce a single value that can be used to compare complex items that are normally difficult to compare (Guba & Lincoln, 1981). PS profile data will situate student achievement data within the varied intellectual contexts created by AAs.

To accurately appraise PS assignments, this index clearly evaluates the transferrable aspects of PS while acknowledging that authentic PS activities are rooted in context. Specifically, the index assesses inherent problem qualities (structure, complexity, abstractness) and the PS environment (clarity of expectations, cues/feedback) to obtain a robust measurement of the challenge posed by the AA. Because student experience largely influences the impact of an assignment, the PSP tool cannot be used solely by an evaluator external to the course; many elements of the index are highly contextual and can be accurately appraised only in consultation with the course instructors. The PSP tool used in this study was validated by two focus groups of educational professionals before it was used by course instructors to assign PS scores to AAs.

## Methods

The methods described here have been approved by the University of Guelph Research Ethics Board (18-05-009). Participants contributed to the study of their own free will and written evidence of informed consent were collected before participants became involved with the study.

### Selection of Problem Solving Assignments

Data was collected from one assignment from each year of the microbiology degree program. The assignments were selected because they were the subject of informal praise by students and other instructors and represent thoughtful, well-designed assessments that mimic real world activities. These assessments were designed by instructors who are actively engaged in the scholarship of teaching and learning; these

individuals were interested in and available to participate in this study. The assignments were selected from three mandatory courses and one restricted elective for the Microbiology Honours undergraduate program. One assignment was selected from each year of study to simulate longitudinal analyses, allowing us to predict how this type of data can be used if it were collected routinely for several years.

### Development and Validation of the Problem Solving Index Tool

A list of factors contributing to the difficulty of a PS assignment was created based on literature exploring characteristics of problems and assessment design principles (Biggs, 1996; Jonassen, 2000; G. Klein et al., 2003; Lazonder, Hagemans, & de Jong, 2010; Mulder et al., 2010; Quintana et al., 2009; Reiser, 2004; Schmidt et al., 2011; Simon, 1988). The tool was formatted as a scale and encourages the user to consider how five PS aspects are represented in their PS assignment (Appendix A).

Ten local experts on rubrics and instructional design participated in the first validation activity. Participants were introduced to the PSP tool and informed of its purpose and construction. Participants were then asked to use the tool on a sample assignment and its rubric and were given a worksheet with feedback prompts. After identifying challenges and areas for improvement, participants offered suggestions to help improve the PSP tool.

The feedback collected from the first validation activity was used to improve the index. Six local experts participated in the second validation (four of whom previously participated in the first validation). The validation exercise was repeated with minor changes and participants raised only a few, surface level concerns (formatting, slightly

different language, etc.). Minor adjustments were made before the PSP tool was ready for use.

### Using the PS Profile tool to assign AAs a PS score

At the beginning of each semester, I scheduled an initial meeting with participating instructors to discuss the details of the study. After obtaining consent from the instructors, I gave them a survey consisting of the PSP tool, the program-level rubric, and additional questions aimed to prepare the instructors to think deeply about their assignment from a pedagogical point of view. I remained in the room while the instructor completed the survey and answered any interpretation questions they had while completing the survey. I did not co-construct any responses; I acted only as a resource to help them interpret the survey. In cases where multiple instructors managed the assignment, two instructors were present for the survey and found consensus before responding to each question. In some cases, to avoid my own misinterpretation of the survey responses, I contacted instructors two weeks after completing the survey to clarify my understanding of their survey responses. The data was analyzed and visualized using Microsoft Excel.

### Reflection

Collecting the PS Profile data was a pleasant experience; instructors were happy that I wanted to use their assignments as PS milestones and were very cooperative throughout the process. My presence was helpful when the instructors completed the survey; they were able to think out loud and bounce ideas off me.

## Results

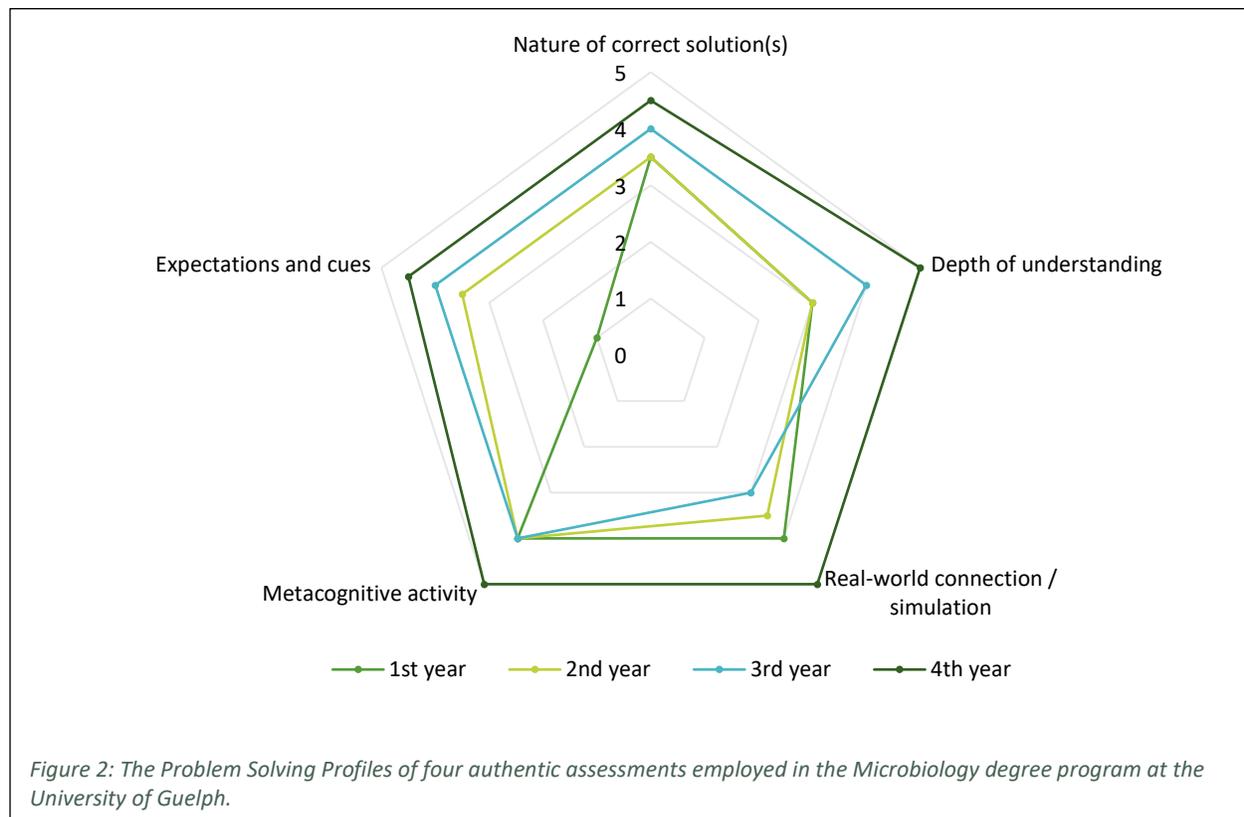
### PS Profiles of four Authentic Assessments

Instructors of the selected AAs used the PSP tool to characterize the PS profile of their AA (Figure 2). The overall PS Profile score (sum of all five PS aspects) increases with year level (Table 3). With the exception of one PS aspect (Real World Connection/Simulation), the PS Profiles indicate maintenance or increased challenge of each PS aspect as year level increases.

*Table 3: Overall PS Profile scores of four authentic assessments employed in the Microbiology degree program at the University of Guelph.*

	Year 1 AA	Year 2 AA	Year 3 AA	Year 4 AA
Overall PS Score*	15.5	17.5	19	24

\*Maximum PS score attainable is 25



## Discussion

### PS Profiles of four Authentic Assessments

These data represent, in part, the diversity of problems and show how different problems elicit thinking in students. Overall PS scores increase with year level, suggesting that more challenging PS assignments are administered as students progress through their degree (Table 3). This finding suggests that the PS complexity of the first, second, and third year assignments are well-scaffolded within the program and that students may have adequate support to help them meet educators' growing expectations until fourth year, when students may face a larger learning gap than in previous years.

The PS Profiles indicate how each of the five PS aspects contribute to the overall challenge of each AA. Most aspects (nature of correct solution, depth of understanding, metacognitive activity, and expectations/cues) either maintain or increase their contribution to the PS challenge with each year (Figure 2). One aspect (real-world connection/simulation) decreases from Year 1 to Year 3 before reaching its maximum value in Year 4. This discrepancy may be an artifact of the study; the four AAs selected were chosen because they are well-designed capstone assessments. In particular, the Year 1 AA exposes first year students to various perspectives within biology and supports the development of numerous soft skills pertinent to success as a scientist. These qualities may explain why the Year 1 AA scored highly in the Real-World Connection / Simulation and Metacognitive Activity aspects.

The 'expectations and cues' aspect increased every year, with its lowest value in Year 1 and its highest value in Year 4. After a large increase between Year 1 and Year

2 (difference of 2.5 units), the 'expectations and cues' aspect experienced small increases between Year 2 and Year 3 (difference of 0.5 units) and Year 3 and Year 4 (difference of 0.5 units). Providing students with enough instruction and examples is an appropriate practice for first year students because these individuals may need additional support as they are introduced to a new discipline (Goff et al., 2015; van Merriënboer et al., 2003). However, too much support hinders the learning process; the gradual removal of support is essential to support students' continued learning and development of self-regulation skills (Ambrose et al., 2010; van Merriënboer et al., 2003). There are currently no best practices to suggest how quickly these supports can be scaffolded, thus giving no indication concerning the appropriateness of the large difference between Year 1 and Year 2.

In summary, these PS Profiles suggest that these four capstone assignments are well-scaffolded overall within the Microbiology program.

### Advantages and Limitations

Although this study exclusively surveyed authentic assessments, the PSP tool is designed to accommodate non-authentic assessments as well. Used extensively throughout a program, PS Profiles may provide a reasonable metric through which educators may compare different assignments and obtain a comprehensive understanding of the program's assessment strategy over a four-year degree.

One limitation of this approach is that the PS Profiles do not indicate students' abilities to solve problems. The PS Profile data must be cross-correlated with student achievement data to provide detailed information about the students' ability to navigate

the problem while considering the situational nuances of the assignment. In other words, when used to gauge learning outcomes achievement, the PSP tool may prove most useful by contextualizing the analysis of other types of student data.

Finally, this study represents the pilot use of the PSP tool. The PSP tool has been validated to the point of first use and, therefore, has not yet been validated by statistical methods used to quantify the validity and reliability of its use. Although such testing may be an area for future work, I emphasize an underlying principle and the original intent of the PSP tool: to leverage the individuality of each AA represented by the uniqueness of each class and instructor who employs the AA. The PSP tool is a flexible tool not meant to be applied by those unfamiliar with the course context; I maintain that each instructor is the expert of their own class and assignment. Such individual application of the tool makes identifying appropriate statistical tests to apply to the tool challenging. To my knowledge, my PSP tool is the only instrument that analyzes PS assignments. It will be interesting for future studies to validate the PSP tool against other tools when they are developed to gauge the difficulty and complexity of PS assignments (American Educational Research Association, American Psychological Association, National Council on Measurement in Education, & Joint Committee on Standards for Educational and Psychological Testing (U.S.), 2014).

## Conclusion

This thesis makes a major contribution to the field of educational development with the first description and piloting of a tool developed to compare the PS content of different assignments. The PS Profile tool is based on cognitive literature and best

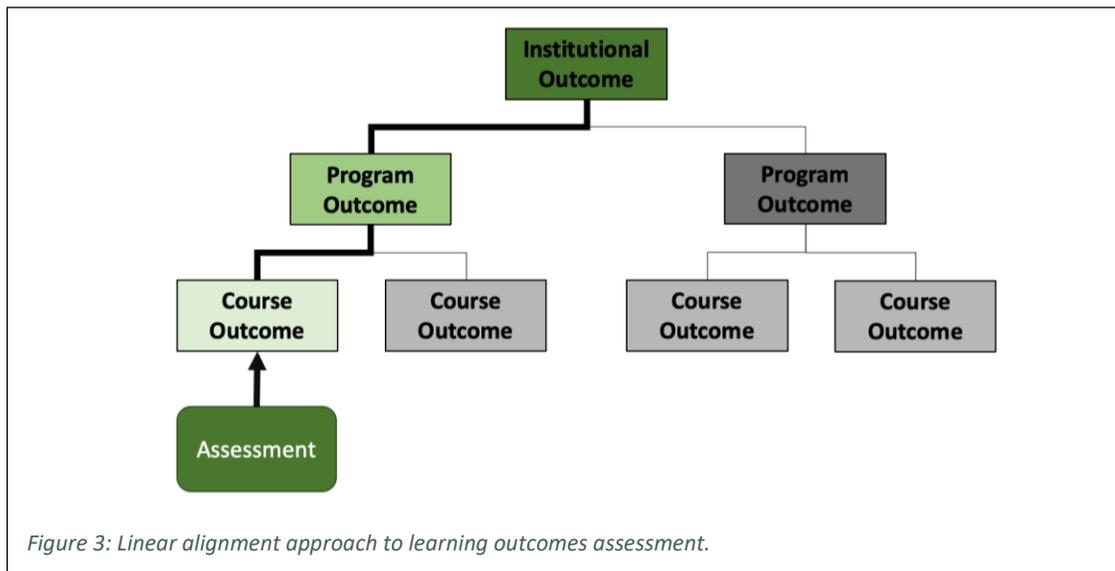
practices concerning assessment design, and provides insights concerning the longitudinal scaffolding of assessments throughout a program. The PS Profile tool is critical in representing the intention perspective foundational to the current thesis methodology. Future LOA projects wishing to include the intention perspective ought to consider implementing the PS Profile tool or using its underlying principles to develop an appropriate variation of this tool.

The data generated by the PS Profile tool can be leveraged in departmental discussions regarding curriculum mapping, learning outcomes assessment, and incorporation of authentic learning experiences (Goff et al., 2015). Although the PS Profile data alone can provoke interesting conversation about program design, coordination with student achievement data (Chapter 3) and student perceptions of assessments (Chapter 4) can provide educators with more meaningful and actionable information.

## Chapter 3: Student Achievement

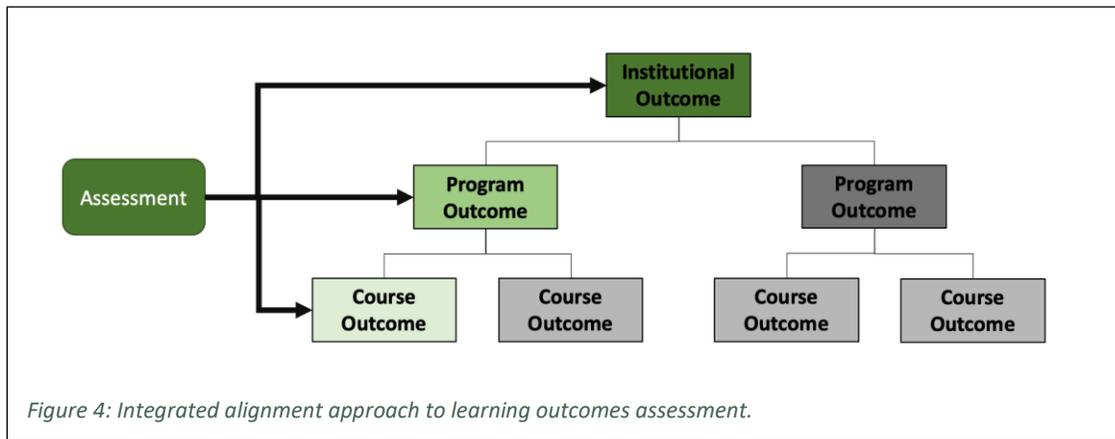
### Learning Outcomes Assessment Strategies

The learning outcomes movement has gained traction in Ontario's post-secondary landscape, particularly at the course and program levels (Goff et al., 2015). Commonly, outcomes assessment projects take the form of curriculum mapping projects where course outcomes are mapped to program outcomes. Additionally, educators often focus on how course outcomes are mapped to assessment activities and how these outcomes are evaluated using task-specific rubrics. These three levels of outcomes (assessment, course, program) exist in a hierarchical relationship (Figure 3).



This framework enables educators to use students' assessment performance to make inferences about their achievement of course and program outcomes (Worthington et al., 2017). While this "linear alignment" approach is quick and intuitive, it is limited in its ability to document nuanced information about complex cognitive skills and students' competencies. The term "nuanced information" is used to describe

detailed information embedded in student work that are indicators of student achievement but may not be represented formally or explicitly in the course learning outcomes or program learning outcomes. Linear mapping of course learning outcomes occasionally results in larger, holistic program learning outcomes being deconstructed into component skills, sometimes oversimplifying the program learning outcome. Many high-quality assignments are holistic assessment activities that may align very closely with a whole program learning outcome. In the linear approach to outcomes assessment, researchers deconstruct the holistic assignment when aligning it to course outcomes and then consolidate a number of course outcomes to align the course to a holistic/sophisticated program outcome. Nuanced information about the students' performances and skills may be lost in the process of deconstructing and consolidating student work. To make the best approximation of program outcomes achievement, key assignments can simply be compared in their wholeness to the appropriate program outcomes in their wholeness in an "integrated alignment" approach to outcomes assessment (an idea I introduce here but that is discussed elsewhere, ex Hodgson, Varsavsky, & Matthews, 2014) (Figure 4). In addition, using an integration strategy to shape curriculum mapping projects ensures that updates to course learning outcomes do not interfere with the accurate mapping of assignments to program learning outcomes; modified course learning outcomes do not affect how an assignment aligns with or accomplishes program outcomes.



This “integrated alignment” strategy of using program outcomes to situate assignments and guide the program can help students identify the transferability of particular skills. This effect can be enhanced with the introduction of a program-level rubric describing the overarching skill that is developed throughout their program. If the students are told what a program outcome looks like solely in the context of one particular assignment, they may struggle to understand how transferrable the skill really is. If students are to identify transferable thinking, educators need to use metacognitive / transferrable language to discuss the skills being taught.

Rubrics are among the most versatile assessment tools available to educators and can be used to help students learn about the skill and discern between accomplished/achieved and developing phases (Massa & Kasimatis, 2018). To support students’ learning and contextualize educational analyses, I suggest the development of a program-level rubric describing the Problem Solving and Critical Thinking program outcome.

## Program Level Rubrics as an Outcomes Assessment Tool

To accurately examine one learning outcome over an entire program, one must mitigate inherent differences between assignments. In this project I examine one assignment from each year of the MICR program, and although they are all part of the MICR calendar, there are too many differences to justify a direct comparison (e.g. differences between assignments, courses, students, and instructors). The program-level rubric is a tool designed to normalize student achievement on the PS aspects of assignments. Assignment rubrics are analyzed and points awarded for PS elements/activities (i.e. concept maps, argumentation, etc.) and are aligned to specific rows on the PS rubric. This alignment helps analysts realize how well students performed on PS activities versus other activities in the assignment (i.e. spelling/grammar, citations). This PS rubric acts as a common ground upon which to compare students in different courses and years of study.

In addition, this PS rubric helps instructors find an intermediate language and framework to guide their students through development of complex cognitive skills. Assignment rubrics are detailed and align with one particular task – sometimes making it challenging for students to conceptualize their transferrable skills beyond that particular task. Institutional rubrics are abstract – able to be interpreted in every discipline and speaking broadly about broad skills, making it challenging for both instructors and students to relate the rubric to every day work/assignments. A program-level skills rubric provides an intermediate tool that can relate to both the detailed assignment rubric and the abstract institutional rubric. Program-level rubrics provide specific yet flexible interpretations of cognitive skills and help users discern between the

different subskills and levels of competency thereof. In this way, the program-level skills rubric can help guide students through the years-long development of PS skills while helping educators understand how their students need to be supported as they move through the development (untrained-to-expert) paradigm.

### Contextualized Outcomes Analysis

In the absence of a program-level rubric for the Bachelor of Science “Problem Solving and Critical Thinking” learning outcome, I developed a rubric *de novo* for this outcome based in the literature. Since this rubric resides at the program level, it focuses on skills instead of assignments. The creation of a new rubric instead of repurposing a rubric already in existence is a rational decision for a number of reasons:

- **Epistemological positioning.** Many rubrics are built on views of problem solving and critical/creative thinking that are fundamentally different from that used in this study. The underlying epistemology influences the development, wording, interpretation, and application of the rubric.
- **Evaluation focus.** Many developmental rubrics assess the processes in which students engage, not the students’ final work (“Types of Rubrics,” 2019). However, given that most problems in undergraduate courses have a correct solution, an accurate assessment of student learning should also consider the final product. If a student can navigate the problem solving steps but consistently fails to successfully solve the problem, I argue that the student is not a proficient problem solver. To be a successful problem solver, students must navigate the process *and* find an appropriate solution.

- **Alignment.** The VALUE rubrics are used extensively as the basis for institutional learning outcomes (“VALUE,” 2019). While these tools outline essential skills learned in classrooms, the rubrics were not intended for use at the classroom level; rather they were meant to be used to guide institutional learning goals (Grouling, 2017). The program-level PS developmental rubric is a tool intended to complement the institutional rubrics by articulating the development of one skill with concrete and detailed language, facilitating alignment with classroom level rubrics.
- **Reliability.** Each institution has a distinct community and language that is commonly understood across campus. This rubric is designed by a University of Guelph student for University of Guelph courses. Two focus groups ensured that the language used in the rubric reflects the ways our educators discuss higher education and helps ensure that the document is interpreted consistently by different folks from across campus.
- **Authenticity.** Our graduates find careers in varied fields. This rubric is informed and shaped by literature about diverse, real-life problem solvers. This authenticity helps our students realize the versatility of their problem solving skills.
- **User-friendly.** This rubric was designed to make its application easier for individuals without a strong background in cognitive science. The reader should not be trying to decipher the rubric; they should be focusing on integrating/identifying elements of problem solving in their courses.

- **Current.** Rubrics are living documents that should be reviewed and updated periodically to accommodate new information.

For this research study, the program-level problem solving rubric was used in conjunction with the PSP tool (see Chapter 2) to collect information about the impact of the assignment and how students' problem solving skills are truly progressing in the Microbiology program.

## Methods

The methods described here have been approved by the University of Guelph Research Ethics Board (18-05-009). Participants contributed to the study of their own free will and written evidence of informed consent were collected before participants became involved with the study.

### Development and Validation of the Program-Level Problem Solving Rubric

This PS rubric models the development of students' PS skills as they progress from untrained problem solvers to expert problem solvers (Appendix B). In real life, PS requires the application of disciplinary, learned knowledge to a unique and complicated scenario. Although successful PS requires correct disciplinary knowledge, this rubric focuses on the student's ability to recognize and use relevant knowledge in new situations. The rubric focuses on the transferrable aspects of PS because the disciplinary knowledge relevant to PS assessments changes with each assessment while the general phases of PS are consistent. The criteria used in this rubric were identified after reviewing numerous PS theories in the literature (ex. Bransford & Stein, 1984; Bryant, 2006; Greiff, Fischer, Stadler, & Wüstenberg, 2015; Ilevbare, Probert, &

Phaal, 2013; Jonassen, 2000; Klein et al., 2003; Matsuo & Nakahara, 2013; Mayr, Smuc, & Risku, 2010; McLennan & Omodei, 1996). The rubric was validated using the same validation method used with the PSP Tool (Chapter 2).

### Alignment of Program-level PS rubric with Institutional Critical and Creative Thinking Rubric

The program-level PS rubric is designed to help educators gauge students' critical and creative thinking based on their PS performance. The University of Guelph lists "Critical and Creative Thinking" as an institutional outcome, and denotes "Problem Solving" as one manifestation of critical and creative thinking (Desmarais, 2012). Our institutional definition of "Problem Solving" is "a process in which one works through a series of operations to come to a conclusion" (Desmarais, 2012). The program-level PS rubric expands upon this definition and explores performance indicators of different levels of problem solving adequacy.

### Alignment of AA rubric with Program-level PS rubric

Instructors were given one survey that was used to both assign the AA a PSP and align the AA rubric with the program-level PS rubric. The method describing the administration of this survey is described in Chapter 2.

### Collection of Student Performance Data

At the beginning of the F18 and W19 semesters, I visited the classes of participating instructors. I introduced the study and requested signed consent from the

students granting me permission to access their graded rubrics after the end of the term. I did not disclose which assignment was being studied. I collected the signed consent forms and handed them to my supervisor, who stored them in a locked filing cabinet in their office until the semester had concluded.

After the end of the semester, the consent forms were given to a disinterested third party consisting of Anamika Dutta (Work Study student, OpenEd) and Jeni Spencer (Educational Analyst, OpenEd). The third party compiled a list of consenting students and contacted the instructors directly requesting the rubrics of said students. After receiving the rubrics, the third party de-identified the rubrics and gave them to me for analysis.

#### Analysis of Student Performance Data

The student performance data was analyzed in SPSS 25 and visualized in Microsoft Excel. Grades of students who dropped the course or who were accused of academic misconduct were not included. A student grade template in Excel was designed for each of the four courses in this pilot study and consisted of a column for student names, student ID numbers, and column(s) for student grades on the entire assignment or assignment components, according to the nature of the assignment. The third party de-identified the data by deleting the columns containing student names and ID numbers; no additional cleaning or formatting was required. The data was then imported into SPSS 25 for analysis.

## Reflection

The analysis of student grades seemed overwhelming at the beginning but was actually a fairly simple process. Because each assignment is unique, the data formatting and linkages to the program-level rubric (collected in the instructor survey) were unique. The biggest challenge was deciding how best to format the data and which statistical tests can be used on this data, which I did with the advice of Lucia Costanzo (Data Resource Centre Analyst, McLaughlin Library, University of Guelph).

## Results

### Rubric Structure

The Program-level PS rubric structure resembles that of an analytic rubric because it employs both rows that describe different elements of the PS process and columns that describe various levels of achievement for each row (Jonsson & Svingby, 2007; Mertler, 2001). The five rows are derived from the literature and include the four common PS steps identified in the literature: 1) understanding the problem, 2) exploring potential solutions, 3) implementing a selected solution, and 4) reflecting on the experience (see Chapter 2). On the rubric, the second step of PS (exploring potential solutions) is broken further into two steps: 'analyzing context' and 'selecting strategies'. 'Exploring potential solutions' is a large, complex step in the PS process and involves different types of thinking and judgment by the problem solver. Given that this rubric is meant to be used in an instructional setting where individuals may be learning about the PS process formally for the first time, I deconstructed the 'exploring potential solutions' to provide some additional guidance and direction to the user.

The rubric uses four columns to describe different achievement levels of a problem solver: untrained, novice, competent, and proficient. The first three columns follow the typical structure of an analytic rubric: one box for each rubric row. This structure implies that a student's performance on the different rows are independent of each other; that a student may be competent at "analyzing context" but may perform poorly on "implementing strategies". This logic does not apply to the fourth column, proficient, because to truly be a proficient problem solver means that the individual has mastered all of the component steps of the PS process. That is, to solve a problem like an expert, the student must be able to expertly navigate every element of PS and in doing so, weave the PS elements together to yield a cohesive solution. This ability to seamlessly move through the PS process is an important quality of proficient problem solvers because authentic problems often require an individual to visit and revisit different PS steps – complex problems are seldom solved by moving through the PS process in a linear fashion. This seamless quality of problem solvers is represented in the rubric by merging the five boxes in the 'proficient' category. This non-typical rubric structure is not compatible with every AA; depending on their underlying design, some AAs cannot be mapped to the 'proficient' box because these assignments are not designed to promote integrated PS, focusing instead on component PS skills.

### Participation

A portion of each class consented to share their grades on the selected AA with the research team (Table 4). These numbers exclude students who consented to participate but were accused of academic misconduct.

*Table 4: Number of students who consented to share their grades on the selected AA.*

Course	Number of students who completed the course	Number of consenting students	Percent participation
Year 1	496	294	59.3%
Year 2	57	35	61.4%
Year 3	60	30	50.0%
Year 4	31	28	90.3%

### Alignment of the PS Rubric with AA rubrics

The instructors of the selected AAs aligned their AA grading scheme to the Program-level PS Rubric (Figure 5). The first year AA mapped to the discrete “Competent” column but could not be mapped to the “Proficient” column, while the second year AA was mapped to the “Proficient” column but not to the “Competent” column. The third year AA mapped to the integrated “Competent” column and the “Proficient” column. The fourth year AA mapped to the discrete “Competent” column and the “Proficient” column. The “Implementing Strategies” criteria of the program-level PS rubric aligned with only one of the selected AAs.

a) Year 1				
Criteria	Untrained	Novice	Competent	Proficient
Understanding the Problem				
Analyzing Context				
Selecting Strategies				
Implementing Strategies				
Evaluating Approach				

b) Year 2				
Criteria	Untrained	Novice	Competent	Proficient
Understanding the Problem				
Analyzing Context				
Selecting Strategies				
Implementing Strategies				
Evaluating Approach				

c) Year 3*				
Criteria	Untrained	Novice	Competent	Proficient
Understanding the Problem				
Analyzing Context				
Selecting Strategies				
Implementing Strategies				
Evaluating Approach				

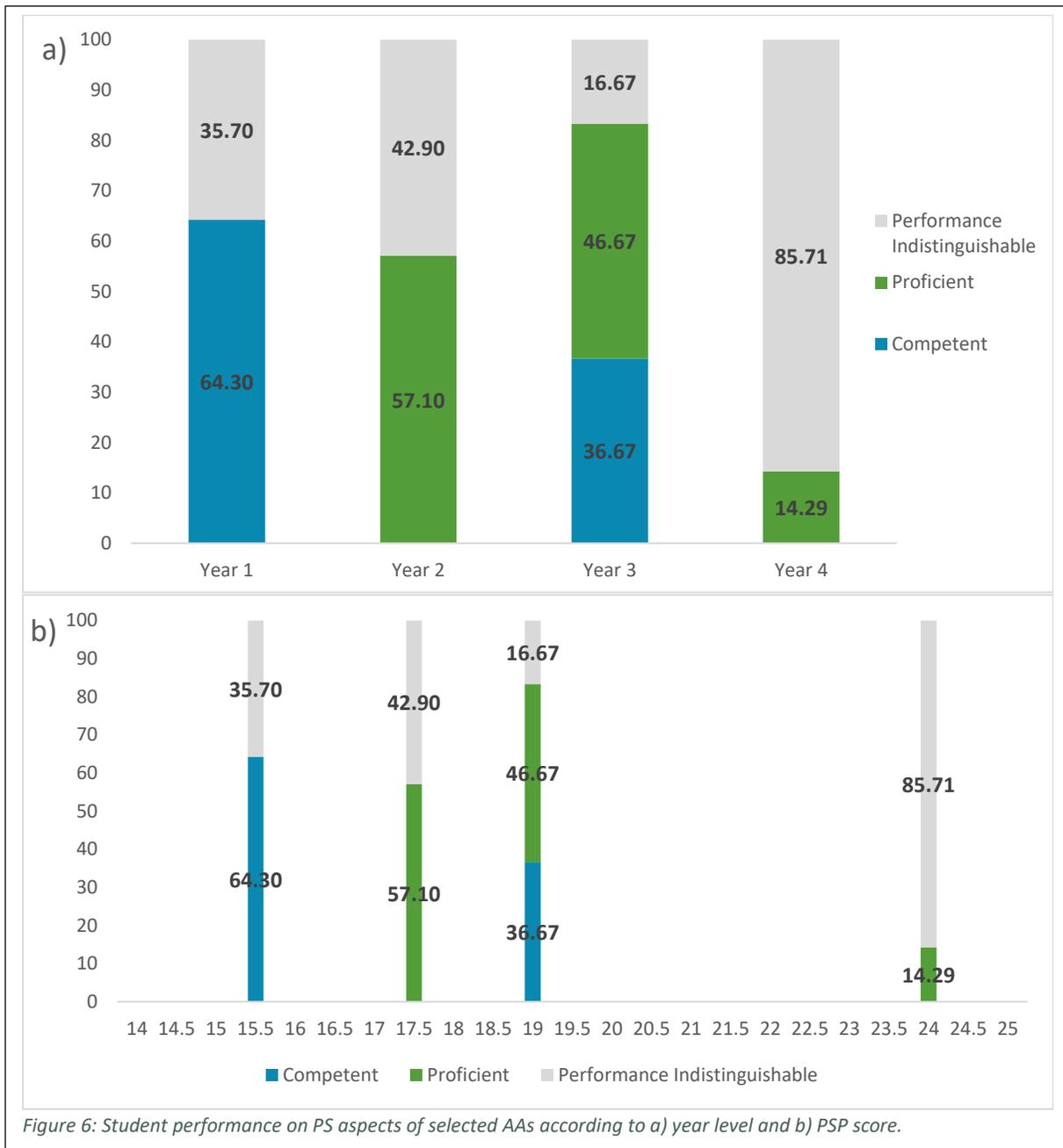
d) Year 4				
Criteria	Untrained	Novice	Competent	Proficient
Understanding the Problem				
Analyzing Context				
Selecting Strategies				
Implementing Strategies				
Evaluating Approach				

\*Due to the nature of the assignment rubric, the Year 3 assignment targets the integration of all five rows on the PS Rubric at both the “Competent” and “Proficient” levels.

*Figure 5: Program-level PS Rubric quadrants targeted by selected assignments of the University of Guelph MICR program.*

## Student Performance

The AA grades of consenting students were analyzed to determine how many students performed at the “Proficient” and/or “Competent” levels (Figure 6). Students whose work did not fully meet the criteria for “Competent” nor “Proficient” achievement fall under the “Performance Indistinguishable” category.



## Discussion

### Alignment of the PS rubric with AA rubrics

The grading schemas of the selected AAs were each successfully aligned with the program-level PS rubric and each alignment was different (Figure 5). This finding

suggests that each of the selected AAs targets students' PS skills in a unique way. These unique linkages may be a property of AAs themselves: a truly authentic assessment is authentic to both the learning activity (real-world connection) and the stage of learning required for the learner (appropriate scaffolding) (Svinicki, 2004). Assignments aiming to support different learning goals for students should have different end points with outcomes that match the level of learning appropriate for the learner and their abilities (Cohen, 1987). The four AAs selected in this study were designed for students at different stages of learning; hence the observation that they align differently to the program-level PS rubric is reasonable and expected.

The Year 1 AA uses a structured, analytical rubric that exclusively targets the Competent column (Figure 5). Perhaps the Year 1 AA introduces students to the various elements of PS without expecting students to demonstrate expert-level PS skills. The Year 2 AA is a semester-long, integrated case study with a grading key that exclusively targets the Proficient column. This alignment suggests that the Year 2 AA encourages students to integrate knowledge and, albeit with introductory-level knowledge, approach complex problems with expert-level skills. The Year 3 AA is a report summarizing a weeks-long independent laboratory experiment designed and performed by each student, with a flexible rubric that aligns with the entire Competent column and the Proficient column. The flexibility of the rubric comes from the graders altering the weight of each row according to how well the student performed on the assignment overall to yield a grade that more accurately reflects the students' work. The alignment with both the entire Competent column and the Proficient column suggests that this AA may be graded for students' ability to report wholly on the laboratory

problem in an integrated manner. The Year 4 AA is a research proposal for an experiment designed to answer a current real-world problem identified by the students, with a holistic rubric that aligns with most Competent quadrants and the Proficient quadrant. The individual alignment with Competent quadrants suggests that the AA provides students with the opportunity to demonstrate their competence at the various elements of PS. The alignment of the Year 4 AA rubric with the Proficient quadrant suggests that students also have the opportunity to demonstrate their ability to complete the task with expert-level skills.

The “Implementing Strategies” criteria of the program-level PS rubric was aligned with only one of the four AAs (Figure 5), because this AA requires students to enact their solutions; the remaining three AAs are simulated experiences that do not provide students with the opportunity to actually implement their proposed solutions.

### Student Performance

Due to the unique linkages between the AA rubrics and the program-level PS rubric (Figure 5), student performance data from each AA must be considered individually and within their context.

The Year 1 AA rubric allowed graders to differentiate between students who achieved competence and those who did not; however, the assignment’s grading rubric could not distinguish between Competent performance and Proficient performance (Figure 5). Of the consenting students, 64.3% achieved Competence, meeting the problem solving standard for the assignment. Since most students were able to perform at the desired performance level, this AA may be an appropriate assignment for first

year students to establish their basic problem solving skills and prepare for the problem solving development to follow in upper years.

The Year AA 2 rubric allowed graders to differentiate between students who achieved Proficiency and those who did not, however the assignment's grading rubric could not distinguish between Competent performance and Proficient performance (Figure 5). Of the consenting students, 57.1% achieved Proficiency, meeting the problem solving standard for the assignment. Over half of students performed at an advanced level of problem solving; however, we do not have sufficient data to suggest how competent the remaining 42.9% of students are with their problem solving skills: perhaps they almost achieved Proficiency, or perhaps their skills were grossly underdeveloped. Additional parameters may be investigated in future studies to further characterize the assignment and the scaffolding leading up to the assignment.

The Year 3 AA rubric allowed graders to differentiate between students who achieved Proficiency, students who achieved Competence, and students whose performance did not meet either of these standards (Figure 5). Of the consenting students, 46.7% achieved Proficiency and 36.7% achieved Competence; in total, 83.4% of students met or surpassed the PS expectation for this assignment. This result suggests that the students are well-prepared to succeed in this assignment and may be a result of effective scaffolding both within the third year course and the program overall. Upper year learning often incorporates skills and knowledge acquired in lower level courses, thus the ability to succeed in complex activities relies, in part, on students' previous learning (Ambrose et al., 2010). At the same time, proper scaffolding

must exist within a single course to support students as they prepare for and engage in more demanding assessment activities (Goff et al., 2015; van Merriënboer et al., 2003).

The Year 4 AA rubric allowed graders to differentiate between students who achieved Proficiency, students who achieved Competence, and students whose performance did not meet either of these standards (Figure 5). 14.3% of students achieved Proficiency while the remaining 85.7% performed at a level that did not meet neither the Competence nor the Proficiency standard. This result suggests that the students may not be prepared to succeed in this assignment, which may be the result of ineffective scaffolding either within the fourth year course or the program overall (van Merriënboer et al., 2003). Analysis of the PSP scores of the four selected assignment suggests that the first year, second year, and third year AAs are scaffolded evenly (an increase of 1.5 – 2 units) whereas the difference between the third year AA and the fourth year AA is larger (an increase of 5 units) (Figure 5). Additional studies may investigate this observed leap from third year to fourth year and identify interventions that may be implemented to further scaffold students' learning. Additional research may identify potential confounding variables that are currently not factored into this study.

### Advantages and Limitations

Although this study saw the development of a rubric targeting only one program-level learning outcome, programs may develop additional rubrics that target other program-level learning outcomes. Aligned with multiple courses, the program-level PS rubric may provide an avenue for faculty members to monitor students' PS achievement from an evaluative, grades-based perspective. Used in conjunction with the PSP tool,

educators can obtain a comprehensive understanding of how students' PS skills evolve over a four-year degree.

One simultaneous advantage and disadvantage of aligning the program-level PS rubric is the unique nature of the alignments with each AA. These customized alignments help convey the uniqueness of each AA and help ensure that the data is not misinterpreted. However, establishing individual linkages with the rubric requires some additional thought when cleaning, formatting, and analyzing the student performance data.

## Conclusion

The program-level PS rubric is a literature-based developmental rubric discussing the differences between untrained and proficient problem solvers and can be used to compare students' PS abilities on completely different assignments. As noted in Chapter 2, to my knowledge there is no literature about how to incrementally scaffold PS skills in assessments across programs, partly because data is not being collected from this perspective. The tools described in this study provide the first tools to study PS across curricula, opening the possibility of studying PS scaffolding across programs.

As with PSP tool data (Chapter 2), program-level PS rubric data can be leveraged in departmental discussions regarding curriculum mapping, learning outcomes assessment, and incorporation of authentic learning experiences (Goff et al., 2015). Combined with the PS Profile data, the program-level PS rubric can provide powerful insights concerning how students actually perform on assignments of varying levels of difficulty. However, the learning cycle is not complete unless the learner is

aware of their own abilities and development (Biggs, 1996). Students should be asked about their perception of their own learning (Chapter 4) to truly understand how our programs function to teach, nurture, and support student development.

## Chapter 4: Student Perceptions

### Significance of the Student Perspective

Increasingly, educational researchers recognize the importance of including the student perspective in educational projects (ex., Biggs, 1996; Brooks, Dobbins, Scott, Rawlinson, & Norman, 2014; Canrinus, Bergem, Klette, & Hammerness, 2017; Cook-Sather, 2006; Jossberger, Brand-Gruwel, van de Wiel, & Boshuizen, 2018; Klemenčič & Chirikov, 2015; P.A. Danaher, 1994; Poth, 2014; Tiessen & Cameron, 2018; Worthington et al., 2017). Some indicators of a constructively aligned curriculum include whether students are capable of the targeted skills *and* are aware that they possess such capabilities (Biggs, 1996). To effectively market their skills to instructors and educators, learners must possess an awareness and understanding of their developed skills (Edge et al., 2018; Martini & Clare, 2014). While the quality of students' skills can be evaluated through assessments, students' awareness of skills must be inferred from student reflections (Biggs, 1996; Tiessen & Cameron, 2018).

### Surveys as Diagnostic Tools

Student surveys are among the most popular data collection tools used in education research (Hackett, 1981; MacFarlane & Brumwell, 2016; Towers, Bannah, & Alderman, 2012; Williams, 2014). Surveys are cost-effective, non-invasive research tools that are tailored to serve a specific purpose (Towers et al., 2012). Researchers often use surveys as screening tools to identify disparities relevant to the study (Klemenčič & Chirikov, 2015) such as gaps between actual and perceived student achievement on assessments or gaps between intended skills gains and awareness of

those gains resulting from completing assignments. Researchers have criticized the extensive use of surveys to collect affective information from students; studies suggest that many surveys are poorly designed, produce invalid datasets, are administered for arbitrary reasons, and the results are not used to inform action (Hackett, 1981; Porter, 2011; Towers et al., 2012; Williams, 2014). However, academics agree that surveys can be valid, reliable, and insightful when designed and applied thoughtfully (Hackett, 1981; Towers et al., 2012; Williams, 2014). Alongside other data analyses, survey responses can provide meaningful insights concerning the impact of a program (Towers et al., 2012; Williams, 2014).

A well-designed survey is an appropriate tool to collect student perceptions in this study. I developed and validated an anonymous, quantitative survey to learn how students perceive their own PS and critical/creative thinking abilities. Cross-referenced with assignment data (Chapter 2) and student performance data (Chapter 3), I investigated whether students feel they are developing PS and critical/creative thinking skills. Overall, students from all year levels are confident in their abilities to communicate both problem solving and critical and creative thinking skills. Comparing the first and second year participants to the third and fourth year participants, students have a significantly different perception of how the selected AA supported the development of their cognitive skills.

## Methods

The methods described here have been approved by the University of Guelph Research Ethics Board (18-05-009). Participants contributed to the study of their own

free will and written evidence of informed consent were collected before participants became involved with the study.

### Development and Validation of the Student Survey

This student survey was designed to collect information about students' perceptions of their own critical and creative thinking skills, problem solving skills, and how these skills are supported by the selected AAs (Appendix C). Students were asked to rate ten items on a five-point Likert scale. The survey questions were designed to investigate the following seven parameters:

- i. whether students relate PS to critical and creative thinking
- ii. students' confidence in their PS skills
- iii. students' confidence in their critical and creative thinking skills
- iv. students' confidence in articulating their PS skills
- v. students' confidence in articulating their critical and creative thinking skills
- vi. students' perceptions of whether the selected AAs support their PS skills
- vii. students' perceptions of whether the selected AAs supports their critical and creative thinking skills

The survey item wording was adapted from a published survey designed to understand student perceptions about learning physics (Adams et al., 2006). Each underlying construct of the survey is represented by two different questions, ensuring that the survey has internal validity and that the questions included in the analysis measure the intended construct (Jhangiani & Chang, 2015). The survey was reviewed and fine-tuned with the input of two staff members of OpenEd (Jennifer Spencer and Aron Fazekas, Educational Analysts) before it was administered to students.

### Administration of the Survey

At the end of the semester, I visited the classes of participating instructors. I reminded the students of the study and revealed which assignment was included in the project. I introduced the students to the anonymous survey and handed out paper copies of the surveys and scantron cards for their responses. The collected scantron cards were submitted to Computing & Communications Services at the University of Guelph for scantron reading. This protocol was used for the first, second, and third year courses. The due date for the fourth year AA occurred after the last lecture, prohibiting a face-to-face administration of the anonymous survey after students submit the AA. To protect the integrity of the survey data (by ensuring that participants had already completed the AA before participating in the survey), I administered the anonymous student survey to the fourth year students via an online platform, Qualtrics. The survey was open for one week and was password protected. I visited the fourth year students during their last lecture, introducing them to the anonymous survey and asking them to respond to the survey invitation that was sent out the following week. The response rates for the surveys are presented in Table 5.

*Table 5: Student participation in the anonymous survey*

Course	Number of students who completed the course	Number of students who participated in survey	Percent participation
Year 1	496	236	47.6%
Year 2	57	53	93.0%
Year 3	60	34	56.7%
Year 4	31	15	48.4%

### Analysis Techniques

After receiving raw scantron readings from Computing & Communications Services, I imported the data into a spreadsheet and formatted the data for analysis using Microsoft Excel and SPSS 25. The fourth-year survey data was downloaded from Qualtrics and formatted to match the scantron data.

Each student's response was analyzed to confirm reliability of the data. Since each parameter of the study was linked to two questions in the survey (Appendix C), I used Spearman's Rank Order correlation to calculate the monotonicity of the relationship between the responses of each pair of linked questions. I chose this test based on what my data permitted without violating any assumptions of the test. Spearman's Rank Order correlation is a nonparametric test used to determine the strength and direction of the relationship between two ordinal variables and has no independence assumptions (Laerd Statistics, 2018b). Spearman's Rank Order correlation is the non-parametric alternative to Pearson's Product Moment correlation (Mukaka, 2012). Although the students in each class are independent persons, statistically they are not independent because they were in the same class. My survey questions asked students to rate each item on a five-point Likert scale, yielding ordinal

data that does not follow a normal distribution. After determining that the correlations are statistically significant, I created bubble charts to visualize the distribution of students' responses to the paired survey questions.

Survey responses from students enrolled in different courses were analyzed to determine whether the student perceptions change as individuals progress through their degree. I used the Kruskal-Wallis H-test to identify statistically significant differences in question responses between students of different year levels. The Kruskal-Wallis H-test is a nonparametric alternative to one-way ANOVA and is used to detect differences in ordinal variables between two or more independent groups (Laerd Statistics, 2018a). The Kruskal-Wallis H-test assumes independence between observations and assumes that the distribution of responses is similar for each group. In my study, each year level is considered an independent group, making each observation (each student response) independent. The distribution of survey data from each group (i.e. class) is similar.

### Reflection

I designed this survey to make the analysis as clear as possible; I wrote questions plainly to minimize misinterpretation of the questions (which would reduce the quality of the data). Additionally, I wrote questions in pairs to add internal validity to the survey. Deciding which questions could be compared was the simplest decision, the more difficult decision was deciding which statistical tests could be used; Lucia Costanzo (Data Resource Centre Analyst, McLaughlin Library, University of Guelph) assisted me in making this decision.

## Results

### Participation

A portion of each class chose to participate in the student survey (Table 5). The Year 1, Year 2, and Year 3 classes were surveyed in class while the Year 4 class was surveyed online. The Year 2 students were surveyed in class before a scheduled quiz, perhaps explaining the higher response rate for these students.

### Validity

Each student's responses to the paired questions were analyzed using Spearman's Rank Order correlation to determine the validity of the responses. Although interpretations vary, correlations larger than  $\pm 0.3$  are often considered significant ("The correlation," 2018). Most responses to each set of paired questions showed a weak (-0.2 to -0.4) or moderate (-0.4 to -0.8) correlation ( $p < 0.01$ ) (Table 6). The negative correlations reflect an inverse relationship between the answers of the paired questions, which is expected given that one question in each pair was framed positively and the other question was framed negatively.

Table 6: Spearman's Rank Order Correlation for paired survey questions.

		Year 1	Year 2	Year 3	Year 4
Paired Questions (to determine data validity)	Comm PS + Trouble Comm PS	-0.457**	-0.384**	-0.500**	-0.570*
	AA helped PS + No change in PS	-0.511**	-0.670**	-0.559**	-0.864**
	Comm CCT + Trouble Comm CCT	-0.468**	-0.386**	-0.463**	-0.453
	AA helped CCT + no change in CCT	-0.596**	-0.695**	-0.604**	-0.659**
Equivalent Questions (to determine how students relate PS to CCT)	Ability to PS + Ability to CCT	0.563**	0.617**	0.440**	0.545*
	Comm PS + Comm CCT	0.790**	0.772**	0.743**	0.907**
	Trouble Comm PS + Trouble Comm CCT	0.641**	0.677**	0.624**	0.710**
	AA helps PS + AA helped CCT	0.672**	0.713**	0.642**	0.611*
	No change in PS + No change in CCT	0.723*	0.819*	0.620**	0.832**

\* significant at  $p < 0.05$  level\*\* significant at  $p < 0.01$  level

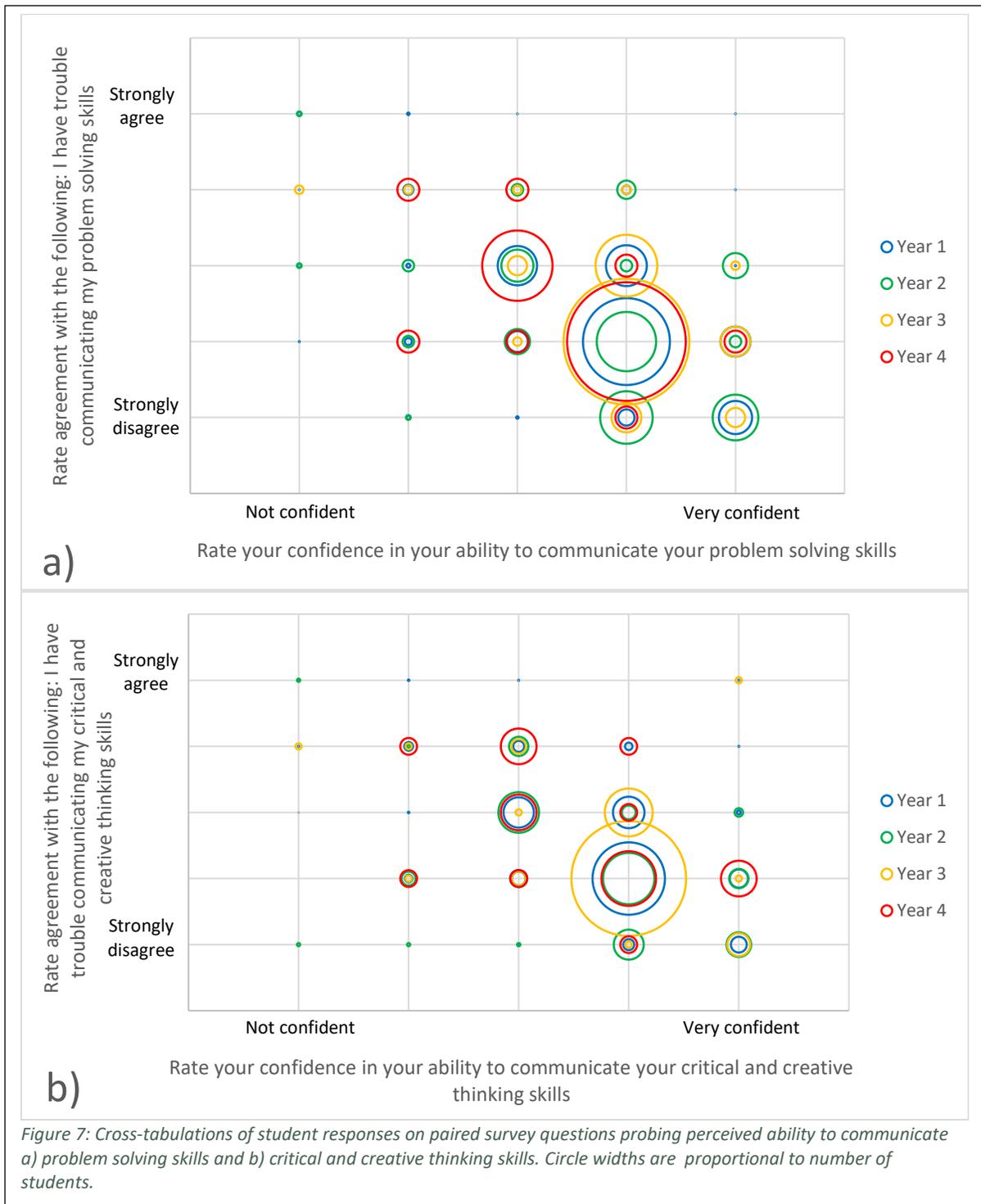
### Association of PS with critical and creative thinking

The student survey consisted of ten questions, with five questions targeting perceived development and ability to communicate PS skills and five equivalent questions targeting critical and creative thinking skills (equivalent questions) (Appendix C). Almost all Spearman correlations between equivalent questions are moderate-strong and positive, suggesting the existence of a positive correlation ( $p < 0.05$ ) (Table 6).

### Perceived ability to communicate PS and critical/creative thinking skills

Student responses to two sets of paired questions (probing students' perceived ability to communicate their PS skills and critical/creative thinking skills) were cross-tabulated and visualized in two bubble charts (Figure 7). The Spearman correlation is significant and negative for the question pair targeting students' perceived ability to

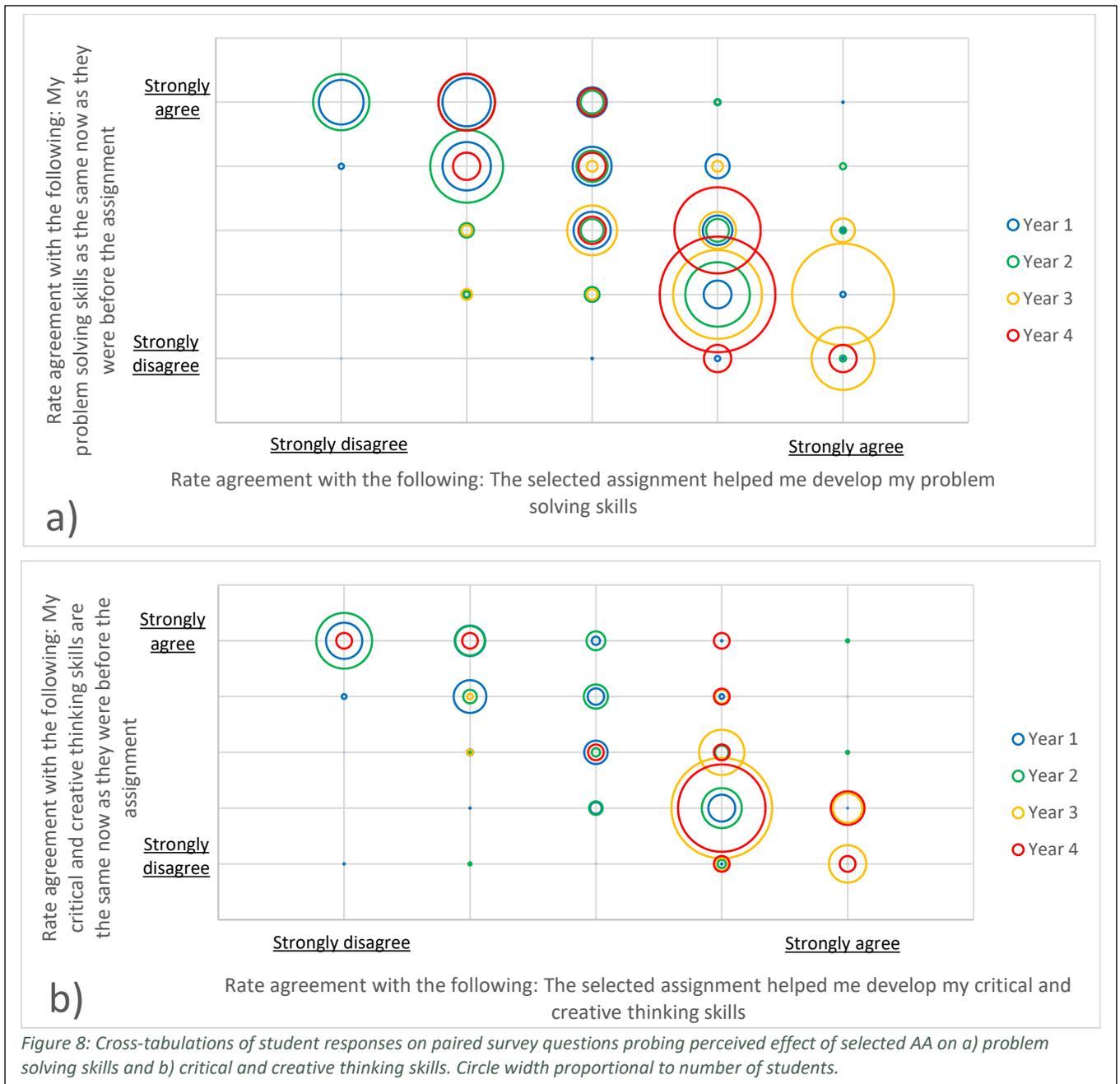
communicate their PS skills (Table 6) and the majority of students indicated that they were confident in communicating their PS skills and that they did not have trouble communicating their PS skills to others (Figure 7). The same result was found for the question pair targeting students' perceived ability to communicate their critical and creative thinking skills.



### Perceived effect of selected AAs

Student responses to two sets of paired questions probing perceived effect of selected AAs were cross-tabulated and visualized in two bubble charts (Figure 8).

Although the Spearman correlation is significant and negative for the question pair targeting effect of AA on problem solving development (Table 6), the majority of first- and second-year students indicated that the selected AA did not help them develop problem solving skills nor did their problem solving skills change after completing the AA. However, the majority of third- and fourth-year students indicated that the selected AA did help them develop problem solving skills and that their problem solving skills are different now than they were before the assignment. The same results were found for the question set targeting effect of AA on critical and creative thinking development.



### Student perceptions over time

Survey responses from students enrolled in different courses were analyzed using the Kruskal-Wallis H-test to determine whether the student perceptions change at different stages of an undergraduate degree (Table 7). Most questions (1, 2, 3, 4, 6, 8) showed no statistical difference across the year levels.

When asked whether the selected AA helped students develop their problem solving / critical and creative thinking skills (questions 5 and 9), Year 3 students answered differently than first- and second-year students, indicating greater agreement that the selected AA helped them develop problem solving / critical and creative thinking skills. When asked whether the selected AA helped students develop their critical and creative thinking skills, Year 4 students also indicated greater agreement to the idea that the selected AA helped them develop critical and creative thinking skills when compared to their first- and second-year counterparts; however, this difference was not observed with the question asking if the AA helped students develop their problem solving skills.

When asked whether the students problem solving / critical and creative thinking skills were the same as they were before completing the AA (questions 7 and 10), Year 3 students answered differently than first- and second-year students, indicating that students felt their skills had changed since completing the assignment.

Table 7: Kruskal-Wallis H-test values for survey question responses across year levels.

Null Hypothesis	Significance	Conclusion
The distribution of Q1 is the same across year level.	0.186	There is no difference in student response based on year level
The distribution of Q2 is the same across year level.	0.206	There is no difference in student response based on year level
The distribution of Q3 is the same across year level.	0.370	There is no difference in student response based on year level
The distribution of Q4 is the same across year level.	0.667	There is no difference in student response based on year level
The distribution of <b>Q5</b> is the same across year level.	<b>0.000</b>	<b>There is a difference</b> in student response based on year level
The distribution of Q6 is the same across year level.	0.216	There is no difference in student response based on year level
The distribution of <b>Q7</b> is the same across year level.	<b>0.000</b>	<b>There is a difference</b> in student response based on year level
The distribution of Q8 is the same across year level.	0.658	There is no difference in student response based on year level
The distribution of <b>Q9</b> is the same across year level.	<b>0.000</b>	<b>There is a difference</b> in student response based on year level
The distribution of <b>Q10</b> is the same across year level.	<b>0.000</b>	<b>There is a difference</b> in student response based on year level

## Discussion

### Validity and Association of PS with critical and creative thinking

Significant Spearman's Rank Order correlations from each year level suggest that questions used in the survey measured the constructs they were intended to measure. Most Spearman correlations between pairs of equivalent questions targeting PS and critical/creative thinking are moderate and positive ( $p < 0.01$ ) (Table 6). This finding suggests that students may associate PS with critical and creative thinking and may be aware that they are deeply connected skills. I did not investigate students' conceptualizations of PS and critical/creative thinking. Further studies may probe how students conceptualize the relationship between PS and critical/creative thinking and how their understanding of these skills evolves as they advance in their degrees.

### Perceived ability to communicate PS and critical/creative thinking skills

Survey responses indicate that most students, regardless of year level, are confident in their ability to communicate their PS skills and that they do not have perceived trouble discussing their PS skills with others (Figure 7). A similar result was found for students' perceived ability to communicate their critical and creative thinking skills. I did not investigate the actual ability of students to communicate their PS and critical/creative thinking skills. However, the observation that students feel confident in their ability to describe these skills to others is interesting, given that other researchers have found that graduates of post-secondary education struggle to communicate their transferrable thinking skills (Edge et al., 2018; Herbert, 2015; Martini & Clare, 2014). This anomaly may indicate that the surveyed students may overestimate their ability to discuss their PS and critical/creative thinking skills; however, further research is needed to fully develop this hypothesis. Additionally, students may not have an advanced understanding of these cognitive skills, thus limiting their ability to accurately discuss their problem solving and critical and creative thinking skills.

### Perceived effect of selected AAs

Survey responses indicate that most first and second year students did not perceive the selected AA as an activity that enhanced their problem solving skills, nor did they perceive any development of their own problem solving skills after completing the assignment (Figure 8, Table 7). However, third-year students responded significantly differently, indicating that they did perceive the selected AA as an activity

that advanced their problem solving skills. A similar result was found for perceived effect of the selected AA on critical and creative thinking skills.

Findings from Chapter 1 suggest that the selected AAs are scaffolded and provide an increased challenge to students as they progress through the degree (Figure 2). Further, with the exception of the fourth year AA, findings from Chapter 2 suggest that students maintain/improve their problem solving achievement, even as the overall challenge of the problem solving assignments increases. Altogether, these data suggest that students are developing their problem solving skills but do not attribute this development to the selected AAs. Students may hold a narrow definition of (or even lack a definition of) the terms “problem solving” and “critical and creative thinking”, which may limit their interpretation of how activities nurture those skills (Deller et al., 2015; Martini & Clare, 2014).

In contrast, perhaps the observed development of problem solving and critical/creative thinking skills can be attributed to the fact that students mature and age as they complete their undergraduate studies (Bhagat, Haque, Bakar, Husain, & Khairi, 2016; Macpherson, 2002). Assuming a relationship between age and maturity, Macpherson found that mature students (aged 30+ years) diverge from dualistic thinking patterns more easily than their less mature counterparts (29 years and younger) (Macpherson, 2002). Assuming a relationship between year level and maturity, the researchers found that mature students (third year) exhibit the same patterns of thought as their less mature counterparts (first year) (Macpherson, 2002). These results suggest that the typical student beginning university at 18 years old may not mature enough over a four year period to have an observable impact on their cognitive habits.

Another facet of maturity includes emotional maturity, which refers to an enhanced self-awareness that is affected by one's intuition, feelings, and spirit (Bhagat et al., 2016). One study found a statistically significant relationship between emotional maturity and the grades of medical students (Bhagat et al., 2016). Beginning students adjusting to medical school displayed signs of poor emotional maturity and experienced lower grades compared to their older counterparts, who displayed signs of both improved emotional maturity and higher grades and learning gains (Bhagat et al., 2016). As students become more emotionally mature, they depend less on the instructor to learn and begin to self-regulate, a crucial skill for learners to establish sophisticated cognitive skills (Bhagat et al., 2016; Macpherson, 2002). Although the surveyed students received no known training about PS and critical and creative thinking, the slightly more mature students in upper years may have intuitively known that the AA had a positive effect on their cognitive skills development.

### Student Perceptions over time

The Kruskal-Wallis H-test found a difference in student responses to questions asking specifically about the selected AA (questions 5, 7, 9, 10) (Table 7). Third-year student responses tended to indicate a greater recognition of the selected AA as an assignment that helped them develop problem solving and critical/creative thinking skills. Interestingly, fourth-year student responses showed no significant difference from first- and second-year students with the exception of one question: fourth-year student responses to question 9 suggested that fourth year students agreed that the selected AA supported the development of their critical and creative thinking skills.

This finding may provide insights about how students relate problem solving to critical and creative thinking. The third-year AA was an individual hands-on laboratory experiment designed and run by students while the fourth-year AA was a written research proposal about a current issue in contemporary microbiology. The third-year students recognized the impact the AA had on both their problem solving and critical and creative thinking skills, while the fourth-year students only recognized the impact of the AA on their critical and creative thinking skills. This subtle differentiation may suggest that students associate problem solving with hands-on, “real life” activities where something has “gone wrong” – more akin to troubleshooting. Problem solving is a broad activity that includes activities that resemble troubleshooting, but extends much farther to include challenges that are more abstract (Jonassen, 2000). With this broader understanding of problem solving, the fourth-year AA can be viewed as a problem solving activity; however, individuals with a narrower definition of problem solving may not recognize how a written proposal has elements of problem solving woven throughout.

A number of variables may also be affecting these results. Perhaps students enter university with the notion that they already have well-refined PS and critical/creative thinking skills. However, as they progress through their degree and their abilities develop, so does their conceptualization of PS and critical/creative thinking. Research on human cognition suggests that students are limited in their ability to generate accurate self-report learning gains, calling into question the validity of any student perception data (Porter, 2011). However, one critical outcome of higher

education is to support students as they learn to become self-aware and communicate the status of their learning, a difficult task to accomplish without consulting students.

## Conclusion

Including the student perspective is crucial for any thorough LOA project; students are experiencing the learning activities we offer and should be aware of both our instructional intentions and their actual learning gains (Biggs, 1996; Edge et al., 2018). The survey administered in this study found that there is no statistical change in students' confidence in their PS and critical and creative thinking skills and that, with the exception of the third-year and sometimes fourth-year students, most students do not view the selected AAs as activities that promote the development of their higher order thinking skills. This is an interesting and important layer of data to consider in conjunction with data describing program expectations (PSP data) and data describing students' actual abilities (student performance data).

## Chapter 5: A More Holistic View of an Undergraduate Science Program

### Considering the Triangulated Data

My thesis research saw the development of a Problem Solving Profile Tool, a program-level developmental problem solving rubric, and a student survey, and the pilot use of all three tools in the Microbiology undergraduate degree program in the 2018-2019 academic year. The coordinated use of these three tools represents the triangulation of three crucial perspectives that ought to be included in LOA projects: program intention, student achievement, and student perception. Considered together, these three types of data provide a richer, more detailed understanding of how an undergraduate program supports students as they develop sophisticated cognitive skills; in this case, problem solving and critical and creative thinking.

Numerous factors contribute to students' abilities to successfully achieve the learning outcomes we define for them and become fully aware of their true capabilities. While many of these factors remain unknown and uncontrolled by educators, strategic interventions or reconfigurations of course/program elements can greatly impact student learning. Collecting data to represent each LOA perspective (intention, achievement, perception) can direct decisionmakers when considering how best to support student learning.

In this pilot study, for example, most students in Years 1, 2, and 3 met the PS standard for their selected AA while the majority of Year 4 students did not (Figure 6). A number of factors could contribute to this result and analyzing additional data can help decisionmakers identify potentially effective interventions. In this scenario, the PS

Profile data suggest that the jump required to move from the Year 3 AA to the Year 4 AA is much larger than the jumps required to move from the Year 1 AA to the Year 2 AA to the Year 3 AA (Figure 6). Educators may consider introducing additional learning modules in Year 3 and Year 4 to help students meet the new learning expectations, or perhaps the curriculum committee may consider re-scaffolding the entire program.

Student perception data provides insight concerning students' awareness of their own skills and how they were supported by the selected AA. Results from this thesis study found no statistical difference between year levels on questions related to student confidence, despite the administration of more challenging PS assignments. Additionally, the only AA that was consistently viewed as one that supports the development of higher cognitive skills was the AA that incorporated hands-on laboratory work. These findings suggest that students' metacognitive awareness of both the nature of higher cognitive skills and the status of their own cognitive skills may be limited. Decisionmakers may consider administering instructional modules that introduce students to metacognition, learning theory, cognitive development models, and explain how the program intends to develop higher order thinking skills through carefully designed assignments.

These data can help direct the efforts of curriculum committees, instructors, and other decisionmakers who are responsible for the development and delivery of undergraduate courses. These data are not meant to be used to value some assessments over others, but rather to characterize the different assessments and recognize the true role each plays in the context of an entire undergraduate degree. As alluded to above, these data may prove useful when designing courses and their

assessments and can guide instructors who wish to explain the intention of their assessments to students and communicate how the assessment is designed to support student development.

### Future Work

This thesis is a pilot study: the first efforts to collect data that triangulates the intention, achievement, and perception perspectives in a LOA project. The methods used within this triangulated methodology include the implementation of a PSP tool, a program-level PS rubric, and an anonymous student survey. These methods were employed on four courses: one from each year of study in the Microbiology program. Future studies may opt to expand the pilot study to include additional courses and assignments. Additionally, future researchers may choose to employ new methods to replace or add to the data collected in this pilot.

Future studies may also move beyond the current methodology and examine how undergraduate students learn about PS and critical and creative thinking, both implicitly and explicitly. Researchers may launch a qualitative study with graduating students to determine which assignments students found most and least helpful in furthering their PS and critical and creative thinking skills. Further research may investigate the knowledge students hold about these cognitive skills at the beginning and end of their undergraduate education and how those conceptualizations change over the course of a degree – if they change at all. Intervention studies may introduce learning modules about metacognition and complex cognitive skills to determine if

explicit training has an impact on students' perceptions of their skills and the role their undergraduate education plays in their development.

### Demonstrating indispensability using data

Detailed, data-driven LOA provide crucial information that can be used to improve post-secondary programs, but can also be used to communicate the true value of university education to prospective students, employers, and the general public. The internet has enabled the upcoming generations of students to grow up with unprecedented levels of communication and access to knowledge. Evolving alongside technology, universities continue to offer valuable learning experiences that help students grow in ways they cannot support on their own – namely in the area of cognitive skills development – including problem solving and critical and creative thinking.

Representatives of higher education ought to use LOA data to communicate the tangible benefits of pursuing tertiary education, and can best do so by describing three perspectives of LOA: the intention of the program, the achievements of enrolled students, and the perceptions of current students and graduates. The world's problems demand solutions attained through superior critical and creative thinking, a skill that is effectively honed at universities. Our future lies in the hands of our prospective students, for they are the next generation of scientists, leaders, and innovators.

## References

- AACU. (2017). Integrating Signature Assignments into the Curriculum and Inspiring Design. Retrieved November 20, 2017, from <https://www.aacu.org/node/4392>
- Adams, W. K., Perkins, K. K., Podolefsky, N. S., Dubson, M., Finkelstein, N. D., & Wieman, C. E. (2006). New instrument for measuring student beliefs about physics and learning physics: The Colorado Learning Attitudes about Science Survey. *Physical Review Special Topics - Physics Education Research*, 2(1), 1–14. <https://doi.org/10.1103/PhysRevSTPER.2.010101>
- Ambrose, S. A., Bridges, M. W., DiPietro, M., Lovett, M. C., & Norman, M. K. (2010). *How learning works: Seven research-based principles for smart teaching* (First). San Francisco: Jossey-Bass.
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education, & Joint Committee on Standards for Educational and Psychological Testing (U.S.). (2014). *Standards for educational and psychological testing*.
- Banta, T. W., Lund, J. P., Black, K. E., & Oblander, F. W. (1996). *Assessment in practice: Putting principles to work on college campuses*. Jossey-Bass.
- Beattie, H. (2012). Amplifying student voice: the missing link in school transformation. *Management in Education*, 26(3), 158–160. <https://doi.org/10.1177/0892020612445700>
- Bhagat, V., Haque, M., Bakar, Y. I. B. A., Husain, R., & Khairi, C. M. (2016). Emotional maturity of medical students impacting their adult learning skills in a newly established public medical school at the east coast of Malaysian Peninsula. *Advances in Medical Education and Practice*, 7, 575–584.
- Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education*, 32(3), 347–364. <https://doi.org/10.1007/BF00138871>
- Blackwell, R. (2018). Is the university experience worth the cost? Retrieved January 30, 2019, from <https://www.theglobeandmail.com/report-on-business/is-the-university-experience-worth-the-cost/article31703109/>
- Bransford, J. D., & Stein, B. S. (1984). *The IDEAL problem solver*. (R. C. Atkinson, G. Lindzey, & R. F. Thompson, Eds.). New York: W.H. Freeman and Company.
- Braun, V., & Clarke, V. (2013). *Successful Qualitative Research: A practical guide for beginners*. (M. Carmichael, Ed.). Washington: SAGE Publications, Inc.
- Brookhart, S. M. (2010). *How to assess higher-order thinking skills in your classroom*. Alexandria, VA: ACSD.
- Brooks, S., Dobbins, K., Scott, J. J. A., Rawlinson, M., & Norman, R. I. (2014). Learning about learning outcomes: The student perspective. *Teaching in Higher Education*, 19(6), 721–733. <https://doi.org/10.1080/13562517.2014.901964>
- Bryant, D. J. (2006). Rethinking OODA: Toward a Modern Cognitive Framework of Command Decision Making. *Military Psychology*, 18(3), 183–206. [https://doi.org/10.1207/s15327876mp1803\\_1](https://doi.org/10.1207/s15327876mp1803_1)
- Bryk, A. S., Nagaoka, J. K., & Newmann, F. M. (2000). *Chicago classroom demands for authentic intellectual work: Trends from 1997-1999*. Chicago, Illinois.
- Campbell, C. M., & Cabrera, A. F. (2011). How Sound Is NSSE?: Investigating the Psychometric Properties of NSSE at a Public, Research-Extensive Institution. *The*

- Review of Higher Education*, 35(1), 77–103. <https://doi.org/10.1353/rhe.2011.0035>
- Canrinus, E. T., Bergem, O. K., Klette, K., & Hammerness, K. (2017). Coherent teacher education programmes: taking a student perspective. *Journal of Curriculum Studies*, 49(3), 313–333. <https://doi.org/10.1080/00220272.2015.1124145>
- Cohen, S. A. (1987). Instructional Alignment: Searching for a Magic Bullet. *Educational Researcher*, 16(8), 16–20. <https://doi.org/10.3102/0013189X016008016>
- Cook-Sather, A. (2006). Sound, presence, and power: “Student Voice” in educational research and reform. *Curriculum Inquiry*, 36(4), 359–390. <https://doi.org/10.1111/j.1467-873X.2006.00363.x>
- Cook-Sather, A., Bovill, C., & Felten, P. (2014). *Engaging students as partners in learning and teaching*. San Francisco: Jossey-Bass.
- Critical Thinking. (2018). Retrieved January 23, 2018, from <https://www.utc.edu/walker-center-teaching-learning/teaching-resources/ct-ps.php>
- Deller, F., Brumwell, S., & Macfarlane, A. (2015). *The Language of Learning Outcomes: Definitions and Assessments*. Toronto, Ontario. Retrieved from [http://www.heqco.ca/SiteCollectionDocuments/The Language of Learning Outcomes-Definitions and Assessments.pdf](http://www.heqco.ca/SiteCollectionDocuments/The%20Language%20of%20Learning%20Outcomes-Definitions%20and%20Assessments.pdf)
- Delpish, A., Holmes, A., Knight-McKenna, M., Mihans, R., Darby, A., King, K., & Felten, P. (2010). Equalizing Voices: Student-Faculty Partnership in Course Design. In C. Werder & M. M. Otis (Eds.), *Engaging Student Voices in the Study of Teaching and Learning* (pp. 96–114). Sterling, VA: Stylus.
- Desmarais, S. (2012). University of Guelph Learning Outcomes. Retrieved March 18, 2016, from [http://www.uoguelph.ca/vpacademic/avpa/outcomes/pdfs/Undergraduate Learning Outcomes.pdf](http://www.uoguelph.ca/vpacademic/avpa/outcomes/pdfs/Undergraduate%20Learning%20Outcomes.pdf)
- Edge, J., Martin, E., & Mckean, M. (2018). Getting to Work: Career Skills Development for Social Sciences and Humanities Graduates, (February).
- Employability Skills. (2018). Retrieved March 27, 2018, from [http://www.conferenceboard.ca/\(X\(1\)S\(hezwd3s304dsx3yusglxz4cm\)\)/spse/employability-skills.aspx?AspxAutoDetectCookieSupport=1](http://www.conferenceboard.ca/(X(1)S(hezwd3s304dsx3yusglxz4cm))/spse/employability-skills.aspx?AspxAutoDetectCookieSupport=1)
- Ewell, P. T. (2009). Assessment, accountability, and improvement: Revisiting the tension. *National Institute for Learning Outcomes Assessment*, (November), 1–24. Retrieved from <http://www.learningoutcomesassessment.org>
- Fielding, M. (2001). Students as Radical Agents of Change. *Journal of Educational Change*, 2(2), 123–141. <https://doi.org/10.1023/A:1017949213447>
- Frederiksen, N., & Ward, W. C. (1978). *Measures for the study of creativity in scientific problem-solving*. *Applied Psychological Measurement* (Vol. 2).
- Gini-Newman, G., & Case, R. (2015). *Creating thinking classrooms*. Vancouver, British Columbia: The Critical Thinking Consortium.
- Goff, L., Potter, M. K., Pierre, E., Carey, T., Gullage, A., Kustra, E., ... Gastel, G. Van. (2015). *Learning Outcomes Assessment: A practitioner’s handbook*. Higher Education Quality Council of Ontario.
- Gordon, J., Ludlum, J., & Hoey, J. J. (2008). Validating NSSE against student outcomes: Are they related? *Research in Higher Education*, 49(1), 19–39. <https://doi.org/10.1007/s11162-007-9061-8>
- Greiff, S., Fischer, A., Stadler, M., & Wüstenberg, S. (2015). Assessing complex

- problem-solving skills with multiple complex systems. *Thinking and Reasoning*, 21(3), 356–382. <https://doi.org/10.1080/13546783.2014.989263>
- Grouling, J. (2017). The path to competency-based certification: A look at the LEAP challenge and the VALUE rubric for written communication. *The Journal of Writing Assessment*, 10(1). Retrieved from <http://journalofwritingassessment.org/article.php?article=120>
- Guba, E. G., & Lincoln, Y. S. (1981). *Effective evaluation*. Jossey-Bass.
- Hackett, G. (1981). Survey Research Methods. *The Personnel and Guidance Journal*, 599–604.
- Hannah. (2019). Is university worth it in 2019? You need to read this before applying.
- Herbert, S. (2015). Focus on outcomes, centre on students: Perspectives on evolving Ontario's university funding model. *Ministry of Training, Colleges and Universities*.
- Higher education still worth the money, new research suggests. (2016). Retrieved January 30, 2019, from <https://www.cbc.ca/news/business/university-college-degree-salary-1.3695254>
- Hodgson, Y., Varsavsky, C., & Matthews, K. E. (2014). Assessment and teaching of science skills: whole of programme perceptions of graduating students. *Assessment & Evaluation in Higher Education*, 39(5), 515–530. <https://doi.org/10.1080/02602938.2013.842539>
- Holmes, N. G., Wieman, C. E., & Bonn, D. A. (2015). Teaching critical thinking. *Proceedings of the National Academy of Sciences*, 112(36), 11199–11204. <https://doi.org/10.1073/pnas.1505329112>
- How has technology changed education? (2019). Retrieved January 31, 2019, from <https://online.purdue.edu/ldt/learning-design-technology/resources/how-has-technology-changed-education>
- Ilevbare, I. M., Probert, D., & Phaal, R. (2013). A review of TRIZ, and its benefits and challenges in practice. *Technovation*, 33(2–3), 30–37. <https://doi.org/10.1016/j.technovation.2012.11.003>
- Is university worth the money? (2018). Retrieved January 30, 2019, from <https://www.theweek.co.uk/98084/is-university-worth-the-money>
- Jhangiani, R. S., & Chang, I.-C. A. (2015). Reliability and validity of measurement. In *Research Methods in Psychology* (Second). Retrieved from <https://opentextbc.ca/researchmethods/chapter/reliability-and-validity-of-measurement/>
- Jonassen, D. H. (2000). Toward a Design Theory of Problem Solving. *Educational Technology Research and Development*, 48(4), 63–86.
- Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, 2(2), 130–144. <https://doi.org/10.1016/j.edurev.2007.05.002>
- Joordens, S., Paré, D. E., & Collimore, L.-M. (2014). *Taking Learning Outcomes to the Gym: An Assignment-Based Approach to Developing and Assessing Learning Outcomes*. Toronto, Ontario.
- Jossberger, H., Brand-Gruwel, S., van de Wiel, M. W. J., & Boshuizen, H. (2018). Learning in Workplace Simulations in Vocational Education: a Student Perspective. *Vocations and Learning*, 11(2), 179–204. <https://doi.org/10.1007/s12186-017-9186-7>

- Klein, G., Ross, K. G., Moon, B. M., Klein, D. E., Hoffman, R. R., & Hollnagel, E. (2003). Macrocognition. *IEEE Intelligent Systems*, 18(3), 81–85. <https://doi.org/10.1109/MIS.2003.1200735>
- Klein, Gary, Associates, K., & Ara, D. (2008). Libro Naturalistic Decision Making. *Human Factors*, 50(3), 456–460. <https://doi.org/10.1518/001872008X288385>.
- Klemenčič, M., & Chirikov, I. (2015). How Do We Know How Students Experience Higher Education? On the Use of Student Survey. In A. Curaj, L. Matei, R. Pricopie, J. Salmi, & P. Scott (Eds.), *The European Higher Education Area*. Springer, Cham. <https://doi.org/10.1007/978-3-319-20877-0>
- Laerd Statistics. (2018a). Kruskal-Wallis H Test using SPSS Statistics. Retrieved February 21, 2019, from <https://statistics.laerd.com/spss-tutorials/kruskal-wallis-h-test-using-spss-statistics.php>
- Laerd Statistics. (2018b). Spearman's Rank-Order Correlation using SPSS Statistics. Retrieved February 21, 2019, from <https://statistics.laerd.com/spss-tutorials/spearman-rank-order-correlation-using-spss-statistics.php>
- Lazonder, A. W., Hagemans, M. G., & de Jong, T. (2010). Offering and discovering domain information in simulation-based inquiry learning. *Learning and Instruction*, 20(6), 511–520. <https://doi.org/10.1016/j.learninstruc.2009.08.001>
- MacFarlane, A., & Brumwell, S. (2016). *The Landscape of Learning Outcomes Assessment in Canada*. Toronto, Ontario.
- Macpherson, K. (2002). Problem-solving Ability and Cognitive Maturity in Undergraduate Students. *Assessment & Evaluation in Higher Education*, 27(1). <https://doi.org/10.1080/0260293012010502>
- Maki, P. L. (2010). *Assessing for learning* (Second). Sterling, Virginia: Stylus Publishing.
- Martini, T. S., & Clare, M. (2014). Undergraduates' Understanding of Skill- Based Learning Outcomes: Can e-Portfolios Help?, 1–44.
- Massa, L., & Kasimatis, M. (2018). Rubrics: Lessons from the Field. Retrieved January 23, 2019, from <https://blogs.illinois.edu/view/915/660903>
- Matsuo, M., & Nakahara, J. (2013). The effects of the PDCA cycle and OJT on workplace learning. *International Journal of Human Resource Management*, 24(1), 195–207. <https://doi.org/10.1080/09585192.2012.674961>
- Mayr, E., Smuc, M., & Risku, H. (2010). Many roads lead to Rome. *Proceedings of the 3rd BELIV'10 Workshop on BEyond Time and Errors: Novel EvaLuation Methods for Information Visualization - BELIV '10*, 8–15. <https://doi.org/10.1145/2110192.2110194>
- McAlpine, L., & Asgharb, A. (2010). Enhancing academic climate: Doctoral students as their own developers. *International Journal for Academic Development*, 15(2), 167–178. <https://doi.org/10.1080/13601441003738392>
- McLennan, J., & Omodei, M. M. (1996). The Role of Prepriming in Recognition Primed Decision-Making. *Perceptual and Motor Skills*, 82(3), 1059–1069. <https://doi.org/10.2466/pms.1996.82.3c.1059>
- Mertler, C. a. (2001). Designing Scoring Rubrics for Your Classroom. *Practical Assessment Research Evaluation*, 7(25), 1–10. Retrieved from <http://pareonline.net/getvn.asp?v=7&n=25>
- Mitchell, K., Shkolnik, J., Song, M., Uekawa, K., Murphy, R., Garet, M., & Means, B. (2005). *Rigor , Relevance , and Results : The Quality of Teacher Assignments and*

*Student Work in New and Conventional High Schools.*

- Mukaka, M. (2012). Statistics Corner: A guide to appropriate use of Correlation coefficient in medical research. *Malawi Medical Journal*, 24(3), 69–71. <https://doi.org/10.1016/j.cmpb.2016.01.020>
- Mulder, Y. G., Lazonder, A. W., & de Jong, T. (2010). Finding out how they find it out: An empirical analysis of inquiry learners' need for support. *International Journal of Science Education*, 32(15), 2033–2053. <https://doi.org/10.1080/09500690903289993>
- Mulnix, J. W. (2012). Thinking Critically about Critical Thinking. *Educational Philosophy and Theory*, 44(5), 464–479. <https://doi.org/10.1111/j.1469-5812.2010.00673.x>
- Newmann, F. M., Bryk, A. S., & Nagaoka, J. K. (2001). *Authentic Intellectual Work and Standardized Tests: Conflict or Coexistence?* Chicago, Illinois.
- Newmann, F. M., Lopez, G., & Bryk, A. S. (1998). *The Quality of Intellectual Work in Chicago Schools : A Baseline Report*. Chicago, Illinois.
- Newmaster, S., Lacroix, C. A., & Roosenboom, C. (2006). Authentic learning as a mechanism for learner centredness. *International Journal of Learning*, 13(6), 103–113.
- Newton, D. (2018). Please stop asking whether college is worth it. Retrieved January 30, 2019, from <https://www.forbes.com/sites/dereknewton/2018/12/16/please-stop-asking-whether-college-is-worth-it/#3a815a4330d2>
- O'Loughlin, A., & McFadzean, E. (1999). Toward a holistic theory of strategic problem solving. *Team Performance Management: An International Journal*, 5(3), 103–120. <https://doi.org/10.1108/13527599910279470>
- O'Neill, G., & McMahan, S. (2012). Giving student groups a stronger voice: using participatory research and action (PRA) to initiate change to a curriculum. *Innovations in Education and Teaching International*, 49(2), 161–171. <https://doi.org/10.1080/14703297.2012.677656>
- P.A. Danaher. (1994). Pupil perceptions of the teacher education practicum: The results of two surveys administered in a Melbourne independent secondary school. *Journal of Education for Teaching*, 21, 25–35.
- Paul, R. W. (1993). The Logic of Creative and Critical Thinking. *American Behavioral Scientist*, 37(1), 21–39. <https://doi.org/10.1177/0002764293037001004>
- Pithers, R. T., & Soden, R. (2010). Critical thinking in education : a review. *Educational Research*, 42(3), 237–249. <https://doi.org/10.1080/001318800440579>
- Pollock, H., Macevic, C., Bajkai, A., Myers, V., Ma, D., Murrant, C., ... Brott, A. (2018). Inspired Research and Scholarship in Biology. College of Biological Science. <https://www.uoguelph.ca/cbs/sites/uoguelph.ca.cbs/files/cbs-strategicPlan-2018-AODA-fixed.pdf>
- Porter, S. R. (2011). Do College Student Surveys Have Any Validity? *The Review of Higher Education*, 35(1), 45–76. <https://doi.org/10.1353/rhe.2011.0034>
- Porter, S. R. (2013). Self-Reported Learning Gains: A Theory and Test of College Student Survey Response. *Research in Higher Education*, 54(2), 201–226. <https://doi.org/10.1007/s11162-012-9277-0>
- Poth, C. (2014). What constitutes effective learning experiences in a mixed methods research course? An examination from the student perspective. *International Journal of Multiple Research Approaches*, 8(1), 74–86.

- <https://doi.org/10.5172/mra.2014.8.1.74>
- Puccio, G., Mance, M., & Murdock, M. (2011). *Creative Leadership: skills that drive change*. (L. Shaw, M. Vail, E. Garner, & P. Schroeder, Eds.) (2nd ed.). Los Angeles: SAGE Publications, Inc.
- Quintana, C., Reiser, B. J., Davis, E. a., Krajcik, J., Fretz, E., Duncan, R. G., ... Soloway, E. (2009). A scaffolding design framework for software to support science inquiry. *Journal of the Learning Sciences*, 13(May 2014), 37–41. <https://doi.org/10.1207/s15327809jls1303>
- Reiser, B. J. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. *Journal of the Learning Sciences*, 13(3), 273–304. [https://doi.org/10.1207/s15327809jls1303\\_2](https://doi.org/10.1207/s15327809jls1303_2)
- Ross, K. G., Klein, G. A., Thunholm, P., Schmitt, J. F., & Baxter, H. C. (2004). The recognition-primed decision model. *Military Review*, 6–10.
- Ruppel Shell, E. (2018). College may not be worth it anymore. Retrieved January 30, 2019, from <https://www.nytimes.com/2018/05/16/opinion/college-useful-cost-jobs.html>
- Schmidt, H. G., Rotgans, J. I., & Yew, E. H. J. (2011). The process of problem-based learning: What works and why. *Medical Education*, 45(8), 792–806. <https://doi.org/10.1111/j.1365-2923.2011.04035.x>
- Signature Assignments and Rubrics. (2016). Retrieved October 3, 2017, from <https://www.ced.csulb.edu/offices/assessment-office/creating-rubric>
- Simon, L. (1988). The Papers We Want to Read. *College Teaching*, 36(1), 6–8.
- Simper, N., Frank, B., Scott, J., & Kaupp, J. (2018). *Learning Outcomes Assessment and Program Improvement at Queen's University*. Toronto, Ontario.
- Sodha, S. (2018). Is your university degree barely worth the paper it's written on? Retrieved January 30, 2019, from <https://www.theguardian.com/commentisfree/2018/mar/16/university-degree-measure-benefits-tuition-fees>
- Statistics Canada. (2017). *Census Profile - Age, Sex, Type of Dwelling, Families, Households, Marital Status, Language, Income, Immigration and Ethnocultural Diversity, Housing, Aboriginal Peoples, Education, Labour, Journey to Work, Mobility and Migration, and Language of Work for*. Ottawa, Ontario. Retrieved from <https://www150.statcan.gc.ca/n1/en/catalogue/98-401-X2016055>
- Steffen, W., Rockström, J., Richardson, K., Lenton, T. M., Folke, C., Liverman, D., ... Schellnhuber, H. J. (2018). Trajectories of the Earth System in the Anthropocene. *Proceedings of the National Academy of Sciences of the United States of America*, 115(33), 8252–8259. <https://doi.org/10.1073/pnas.1810141115>
- Svinicki, M. D. (2004). Authentic assessment: Testing in reality. *New Directions for Teaching & Learning*, 100, 23–29. <https://doi.org/10.1002/tl.167>
- The correlation. (2018). Retrieved June 14, 2019, from <https://allpsych.com/researchmethods/correlation/>
- Thiessen, D. (2006). Student knowledge, engagement, and voice in educational reform. *Curriculum Inquiry*, 36(4), 345–358. <https://doi.org/10.1111/j.1467-873X.2006.00362.x>
- Tiessen, R., & Cameron, J. (2018). The Relationship Between Experiential Learning and Career Outcomes for Alumni of International Development Studies Programs in

- Canada. *Canadian Journal of Higher Education*, 48(3), 23–42.
- Towers, S., Bannah, S., & Alderman, L. (2012). Student feedback systems in higher education: a focused literature review and environmental scan. *Quality in Higher Education*, 18(3), 261–280. <https://doi.org/10.1080/13538322.2012.730714>
- Types of Rubrics. (2019). Retrieved January 24, 2019, from <https://resources.depaul.edu/teaching-commons/teaching-guides/feedback-grading/rubrics/Pages/types-of-rubrics.aspx#developmental>
- VALUE. (2019). Retrieved January 24, 2019, from <https://www.aacu.org/value>
- van Merriënboer, J. J. G., Kirschner, P. A., & Kester, L. (2003). Taking the Load Off a Learner's Mind : Instructional Design for Complex Learning Taking the Load Off a Learner's Mind : Instructional Design for Complex Learning. *Educational Psychologist*, 38(1), 37–41. <https://doi.org/10.1207/S15326985EP3801>
- Walsh, B., Ciais, P., Janssens, I. A., Peñuelas, J., Riahi, K., Rydzak, F., ... Obersteiner, M. (2017). Pathways for balancing CO2 emissions and sinks. *Nature Communications*, 8. [https://doi.org/10.1016/0306-9877\(90\)90061-I](https://doi.org/10.1016/0306-9877(90)90061-I)
- Wiggins, G. (1990). The Case for Authentic Assessment. *Practical Assessment, Research & Evaluation*, 2(2), 1–4. <https://doi.org/ED328611>
- Williams, J. (2014). Student feedback on the experience of higher education a significant component of institutional research data. In *Using Data to Improve Higher Education: Research, Policy and Practice* (pp. 67–80). [https://doi.org/10.1007/978-94-6209-794-0\\_5](https://doi.org/10.1007/978-94-6209-794-0_5)
- Wilson, M., & Sloane, K. (2010). From Principles to Practice : An Embedded Assessment System. *Applied Measurement in Education*, 13(2). <https://doi.org/10.1207/S15324818AME1302>
- Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, 17, 89–100. <https://doi.org/10.1111/j.1469-7610.1976.tb00381.x>
- World Economic Forum. (2017). *Summary for Policymakers. Climate Change 2013 - The Physical Science Basis*. <https://doi.org/10.1017/CBO9781107415324.004>
- Worthington, P., Dewancker, A., LaRush, N., Lackeyram, D., & Dawson, J. F. (2017). Engaging student stakeholders in developing a learning outcomes assessment framework. In *Discussions on University Science Teaching: Proceedings of the 2015 Western Conference on Science Education* (Vol. 1). Retrieved from <http://ir.lib.uwo.ca/wcsedust/vol1/iss1/8>

## Appendices

### Appendix A: The Problem Solving Profile Tool

Instructors used the following tool to characterize the problem solving challenge of their AA.

(1) Novice-level problems		Expert-level problems (5)	
<i>Nature of correct solution(s)</i>			
Students achieve success by simply following disciplinary rules  Creative thinking not essential to earning a high grade  <i>Examples: most end-of-chapter textbook questions, drills</i>		Students achieve success by thinking creatively about the discipline / scenario  Substantial amount of creative thinking essential to earning a high grade  <i>Examples: designing an experiment to answer a novel problem, inventing / proposing a new product</i>	
1	2	3	4
5			
<i>Depth of understanding</i>			
Success can be achieved via memorization or surface understanding of the topic  <i>Examples: most end-of-chapter textbook questions, drills</i>		Success requires students to have an advanced understanding of the discipline  <i>Examples: analyzing authentic scientific data (with multiple considerations), predicting outcomes of a complex scenario</i>	
1	2	3	4
5			
<i>Real-world connection / simulation</i>			
Assignment isolates disciplinary concepts from context  Assignment does not require any additional research  <i>Examples: most end-of-chapter textbook questions, drills</i>		Assignment is situated in a real-world / simulated context  Assignment requires students to learn material beyond the course content  <i>Examples: designing a research study on a current topic, using the discipline to solve a problem with political/cultural value, community-based projects</i>	
1	2	3	4
5			
<i>Metacognitive activity</i>			
Assignment does not require students to reflect on their solution or do so superficially  <i>Examples: only asking for the solution, asking students to 'show their work'</i>		Assignment requires students to critically evaluate the effectiveness of their solution and potential implications  <i>Example: asking students to explain what went wrong / what worked, what would be different if they used another solution</i>	
1	2	3	4
5			
<i>Expectations and cues</i>			
Assignment provides step-by-step instructions  <i>Examples: tutorials for the assignment</i>		Assignment requires students to navigate the task on their own.  <i>Examples: students seek their own assistance by booking appointments with resource centers outside of the course</i>	

Please note that ample direction and support for students is appropriate in some contexts, for instance in first or second year as students are learning about the discipline. Also note that as students advance in their degree, clear assignment expectations are still required, but perhaps with less direction about *how* to successfully complete the assignment.

**1****2****5****3****4**

## Appendix B: The Program-Level PS Rubric

Instructors mapped their AA rubric to the following program-level PS rubric:

### Problem Solving Developmental Rubric

Here, a problem is considered as a situation where there is a difference between the current state and the desired state, and the pathway to achieve the desired state is not obvious or prescribed. In real life, problem solving requires a person to use their critical and creative thinking with their disciplinary knowledge to navigate a unique, complicated scenario. Proficient problem solvers seamlessly integrate the different components of problem solving – this is depicted by merging the cells under the ‘Proficient’ category.

Criteria	Untrained	Novice	Competent	Proficient
<b>Understanding the Problem</b> Distinguishes between the current problem state and the desired state	Incorrectly distinguishes current state from desired state	Ambiguously distinguishes current state from desired state	Clearly distinguishes current state from desired state	Clearly distinguishes current state from desired state
<b>Analysing Context</b> Considers factors contributing to the problem and any situational circumstances	Exclusively analyses factors immediately implicated in the problem	Analyses factors implicated in the problem and some situational circumstances.	Analyses most factors involved in the problem solving scenario	Analyses most / all factors in the problem solving scenario. Challenges assumptions and considers multiple perspectives.
<b>Selecting Strategies</b> Chooses strategies to respond to the problem	Strategies do not contribute to solution  Rationale for strategy selection is incorrect or absent	Strategies contribute little to solution  Rationale for strategy selection is superficial	Strategies contribute adequately to solution  Rationale for strategy selection is conceptual	Student selects the best strategies for the problem  Student provides advanced rationale for entire problem solving approach  Efficiently implements strategies
<b>Implementing Strategies</b> Carries out strategies to solve the problem	Incompetently / fails to implement strategies	Implements strategies with challenge	Adequately implements strategies	Reflects on success of their overall response to the problem, specific strategies employed, and contemplates areas for improvement Thoughtfully discusses the interdisciplinary / real-world implications of the problem when appropriate.
<b>Evaluating Approach</b> Considers the effectiveness of their response to the problem	Does not reflect on the effect of their response to the problem  Does not discuss / incorrectly discusses the interdisciplinary / real-world implications of the problem.	Reflects on overall success of their response to the problem  Superficially discusses the interdisciplinary / real-world implications of the problem when appropriate.	Reflects on success of their overall response to the problem and specific strategies employed  Conceptually discusses the interdisciplinary / real-world implications of the problem when appropriate.	

## Appendix C: Student Perceptions Survey

Students were asked to anonymously answer ten questions probing their perceptions of their own PS and critical/creative thinking skills (Figure 9).

Using the provided scantron sheet, please rate your confidence in your ability to:

	Not confident			Very confident	
	A	B	C	D	E
1. Solve real-world biology problems					
2. Think critically and creatively	A	B	C	D	E
3. Communicate my problem solving skills to others	A	B	C	D	E
4. Communicate my critical and creative thinking skills to others	A	B	C	D	E

Using the provided scantron sheet, please indicate your level of agreement with the following statements:

	Strongly disagree			Strongly agree	
	A	B	C	D	E
5. The [selected AA] helped me develop my problem solving skills.	A	B	C	D	E
6. I have trouble describing my critical and creative thinking abilities to others.	A	B	C	D	E
7. My problem solving skills are the same now as they were before completing the [selected AA].	A	B	C	D	E
8. I have trouble describing my problem solving skills to others.	A	B	C	D	E
9. The [selected AA] helped me develop my critical and creative thinking skills.	A	B	C	D	E
10. My critical and creative thinking skills are the same now as they were before completing the [selected AA].	A	B	C	D	E

Figure 9: The anonymous student survey administered in this study.

Each question used in this survey is linked with at least one other question (see below), enabling cross-tabulation of student responses which allows us to make inferences about student perceptions. Paired questions ask students about the same topic but use different language; these paired questions were used to gauge the validity of responses (i.e. determine whether students answered each question at random) (Table 8).

*Table 8: Paired survey questions used to gauge validity of student responses.*

Topic	Survey Questions	Expected Correlation
Perceived ability to communicate PS skills	3 and 8	Strong, negative
Perceived ability to communicate critical and creative thinking skills	4 and 6	Strong, negative
Perceived effect of AA on PS skills	5 and 7	Strong, negative
Perceived effect of AA on critical and creative thinking skills	9 and 10	Strong, negative

If the survey responses are considered valid, the aggregate responses to each question may be analyzed. Each pair of survey questions targets a topic of interest to this study (Table 8). In addition, the ten survey questions were coordinated to probe students' understanding of the relationship between PS and critical/creative thinking (Table 9). Students answered five questions probing their perception of their own PS skills and five equivalent questions probing their perception of their own critical and creative thinking skills. These equivalent questions used the exact same wording, enabling a direct comparison of questions targeting PS and questions targeting critical and creative thinking.

*Table 9: Equivalent survey questions used to compare student perceptions of PS and critical and creative thinking.*

Equivalent PS question	Equivalent critical and creative thinking question
1	2
3	4
5	9
7	10
8	6