The Role of ICAP in Effective Course Design: A Learning Analytic Evaluation

by

Jesse Ha

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2022 by the
Graduate Supervisory Committee:

Brian C. Nelson, Chair
Michelene T. H. Chi
Margarita Pivovarova

ARIZONA STATE UNIVERSITY

May 2022

ABSTRACT

Recently, the Interactive-Constructive-Active-Passive (ICAP) framework has been gaining increasing prominence in cognitive and learning sciences. The ICAP theory asserts that students learn more deeply when they are cognitively engaged in generative and collaborative learning. Indeed, prior studies have established the value of the ICAP framework for predicting student learning. However, the framework has yet to become widely used by practitioners, possibly due to the lack of accessible resources for applying the framework instruction design. This study sought to fill that gap by implementing and validating the ICAP instructional rubric instrument to rate the design of college chemistry courses at a large public university in the southwest and exploring its relationships with several metrics of student performance via multiple regression analysis: a) level of participation; b) final exam grades; c) course grades; d) course retention; and e) course attrition. This study analyzed data from the university's learning management system and included student-level controls such as markers of prior academic performance (i.e., GPA and SAT scores) as well as student demographics. The findings of this study suggest that the ICAP framework may be a useful tool for instructors to improve course design. In addition, the ICAP framework's predictive claims on student deeper learning were further validated by the results of this study.

DEDICATION

This work is dedicated to the love and anchor of my life, Dr. Paula Yoon. Thank you for supporting me in the entirety of the doctoral studies journey that has culminated in this document. You were right in suggesting that I uninstall my video games temporarily so that I could write this paper on time. We made it!

This work is also dedicated to my extended family and particularly in memory of my uncle, Ho Soon Chang, who always took the time to support me with alacrity.

ACKNOWLEDGEMENTS

This work would not have been possible without the guidance of my doctoral committee. I would like to acknowledge Drs. Brian Nelson, Michelene Chi, and Margarita Pivovarova for their support and encouragement throughout my doctoral studies. In addition, I would like to acknowledge Dr. James Gee who took the time to mentor me personally. My skills and abilities as a researcher and colleague have been greatly shaped by their influence on me. I would also like to acknowledge Dr. Philippos Savvides, Dr. James Cunningham, and Ruvi Wijesuriya for their support on this project.

I would also like to acknowledge the incredible friendships and partnerships I have fostered along this journey that have made it more bearable and oftentimes enjoyable, in particular, Man and Danielle for their support throughout our program.

I would like to acknowledge all the colleagues and students who I have crossed paths with throughout my journey as an educator. There are many whom I have had the privilege of instructing, learning from, mentoring, collaborating with, coaching, tutoring, and counseling. These formative experiences and relationships have acted as my internal compass by which I have navigated my career and calling. In particular, I would like to acknowledge Amber, Betty, Brian, Chelsea, Christine, Dave, students of Dream Team, Elena, Juliana, Michelle, Angie, Rachael, and Stormy.

Finally, I would like to acknowledge friends who have all had their share in dealing with my shortcomings and have nonetheless been foundationally supporting me on this journey. In particular, I would like to acknowledge Yera, Eric, Sam, James, Ara, Daniel, Monica, Isaac, Bora, Jihae, Chris, Kristin, Tim, Esther, Pierre, Jua, Linda, Christine, Wanhee, Kimmie, Josh, Livy, Kurt, Clara, Hieyeon, and Charse.

TABLE OF CONTENTS

Page

LIST OF TABLES……………………………………………….………..………..vii

LIST OF FIGURES ……………………………………………………...……...viii

CHAPTER

1 INTRODUCTION ……………………………………...…………………1

Statement of the Problem……………………………………...………1

Research Questions ………………………………………………………4

2 THEORETICAL FRAMEWORK ………………………………………………5

The ICAP Framework …………………………………………………………5

Passive Mode …………………………………………………...………6

Active Mode……………………………………………………….………8

Constructive Mode …………………………………………………...……9

Interactive Mode ………………………………..…………………..……10

Ascertaining the Mode of Engagement……………………..……………10

ICAP Assumptions ……………………………………………….……...11

Other Considerations ……………………………………………….……13

Related Theoretical Perspectives ……………………………………….…14

Behaviorism ……………………………………………………………...…14

Cognitivism……………………………………………………………………17

Constructivism ……………………………………………………...…23

Branches of Constructivism…………………………………………...…25

Cognitive Constructivism ……………………………………………25

CHAPTER Page

Social Constructivism 27

Related Frameworks 28

Bloom's Taxonomy 28

Webb's Depth of Knowledge..........30

3 BACKGROUND LITERATURE..........33

Online Learning 33

LMS and Course Participation..........35

Course Retention..........36

Design and Evaluation of College Courses 37

4 METHODOLOGY 39

Research Questions & Rationale 39

Research Design..........40

Data 40

Instrumentation 41

Data Procedures 45

5 RESULTS & DISCUSSION OF FINDINGS..........46

Descriptive Statistics..........46

Results..........55

Course Participation..........58

Final Exam Grade 65

Course Grade 73

Course Retention..........82

CHAPTER Page

Course Attrition 90

6 LIMITATIONS 93

7 IMPLICATIONS AND RECOMMENDATIONS 96

The ICAP Instructional Rubric 96

Course Design and Evaluation 100

Learning Analytics 102

REFERENCES 104

APPENDIX 117

A THE ICAP INSTRUCTIONAL RUBRIC 117

B INSTITUTIONAL REVIEW BOARD (IRB) APPROVAL LETTERS 130

LIST OF TABLES

Table Page

1. Examples of ICAP Rating ……………………………………………..…………44

2.Overall Descriptive Statistics for Participant Demographics ……...………………...48

3. Overall Descriptives for Dependent Variables ……………………………..…….50

4. Overall Descriptives for Independent Variables…………………………...……..52

5. Course Participation…………………………………………………………...….57

6. Course Participation - Descriptive Statistics - Significantly Related Variables…….61

7. Final Exam Grade ……………………………………………………………64

8. Correlation Matrix for ICAP Dose – Final Exam Grade ……………………..…..66

9. Final Exam Grades - Descriptive Statistics - Significantly Related Variables…...…69

10.Course Grade ……………………………………………………….……………...72

11.Correlation Matrix for ICAP Dose – Course Grade ………………………………..74

12. Course Grade - Descriptive Statistics - Significantly Related Variables……….….77

13. Course Withdrawal ……………………………………...…………...…..81

14. Summary for Course Retention ………………………………..……………….....85

15. Number of Weeks Participated……………………………………….…………...89

16. Course Attrition - Descriptive Statistics - Significantly Related Variables……….92

17. Significant Findings for the ICAP Instructional Rubric ………………………….95

18. Significant Findings for Learning Analytics ……………………….…………………101

LIST OF FIGURES

Figure Page

1. The ICAP Framework ………………………………………………………7

CHAPTER 1

Introduction

Statement of the Problem

Over the past decade, the Chi and Wylie's (2014) Interactive-Constructive-Active-Passive (ICAP) framework has been given much consideration in the fields of cognitive science and the learning sciences. The ICAP framework theorizes that students who are cognitively engaged in generative learning and collaboration are likely to learn more deeply. The ICAP theory was first introduced in Chi's (2009) seminal study, which proposed three cognitive modes of engagement (active, constructive, and interactive), along with evidence in the literature that supported the ICAP framework's predictions for deeper student learning with the following hierarchy: interactive > constructive > active. ICAP was further extended in Chi and Wylie's study to include the passive mode, which accounted for numerous laboratory and classroom studies in the body of literature that showed differences in learning gains between the passive mode and the active mode. The full hierarchy follows the ICAP acronym: interactive > constructive > active > passive.

The ICAP framework is often juxtaposed with Bloom's Taxonomy (BT) and Webb's Depth of Knowledge (DOK). The ICAP framework differs from BT in the sense that it can be used for more than instructional sequencing or assessment/task creation afforded by BT. In comparison to Webb's DOK, the ICAP framework focuses on how students are cognitively engaging rather than on the complexity of task or assessment content. The ICAP framework potentially assists instructors in designing and evaluating instruction at the level of student cognition whereas BT and Webb's DOK are limited to designing and evaluating instruction at the level of task type or complexity. Increasingly, scholars are utilizing the ICAP framework for framing their studies; however, to date, accessible and convenient pedagogical tools and instruments such as the BT pyramid or Webb's DOK wheel for applying the ICAP framework to instruction are lacking.

In addition to the progression of learning and cognitive sciences, instructional technology has developed at a breakneck pace in the past decade. What was once considered a fad or impractical and utilized mainly in distance education, online learning is now ubiquitous. In tandem with online learning, the nascent field of educational data analytics has emerged since learning within the computing space has afforded researchers with heaps of readily stored data on student interactions within learning management systems (LMS).

Recent literature suggests that more research is needed that focuses on LMS software, including learner outcomes such as retention and learner success as well as evaluation and quality assurance (Martin et al., 2020). Smith et al. (2012) found that LMS participation markers have a predictive effect on student course outcome. However, the study, like many others, does not explain the factors that predict why some students participate more in the LMS. More recently, studies suggest that future applications of data analytics should specifically address "active learning" since increased participation predicts increased student performance (Avci & Ergun, 2019; Darko, 2021). In addition, education data analytics also afford researchers with easily accessible data on student course retention and attrition. Prior research has shown that "active learning" methods may improve student retention in college chemistry courses (Shattuck, 2016) and programs of study (Braxton et al., 2000; Van den Berg & Hofman, 2005). Given that the ICAP framework is an extended framework for "active learning", it offers a potential explanation since students engaged in the three higher modes of the ICAP framework utilize their cognitive resources differentially while participating in the LMS via “active learning”.

Given the ICAP framework predicts deeper student learning, these markers of performance and outcome may have relationships with the ways in which instructors design their courses. Accordingly, rating the design of courses using an instrument based on the ICAP framework may elucidate trends of student academic performance, participation, and retention/attrition.

The general problems at hand are a gap in literature regarding the application of the ICAP framework at the practitioner level via accessible tools/instruments and the reasons for why students may be more or less likely to utilize the course LMS. This study addressed these two problems by contributing to the literature in a three-fold manner: a) testing the use of a more accessible instrument, the ICAP instructional rubric, for designing and evaluating instruction; b) exploring the relationship between ICAP ratings of courses and student participation in the LMS; and c) validating the use of the ICAP instructional rubric in course design and evaluation by relating it to various markers of student performance and outcomes.

In this study, I evaluated the quality of course design of postsecondary online and face to face (F2F) introductory chemistry courses using the ICAP Instructional Rubric and assessed its impact on and relationship with related student course outcomes: student participation level, final exam grade, course grade, course retention, and course attrition.

Research Questions (RQ)

1) To what extent are learning activities and learning materials, as measured by the ICAP instructional rubric, associated with student participation in course LMS?

2) To what extent do learning activities and learning materials, as measured by the ICAP instructional rubric, predict student final exam scores?

3) To what extent do learning activities and learning materials, as measured by the ICAP instructional rubric, predict student course grades?

4) To what extent are learning activities and learning materials, as measured by the ICAP instructional rubric, predictive of student course retention?

5) To what extent are learning activities and learning materials, as measured by the ICAP instructional rubric, predictive of student course attrition?

CHAPTER 2

Theoretical Framework

The ICAP Framework

I begin this section by briefly overviewing the ICAP framework and discussing the components and nuances of the ICAP framework. In the following section, I explore the relationship between the ICAP framework and related theoretical perspectives: behaviorism, cognitivism, cognitive constructivism, and social constructivism as well as the relationship between the ICAP framework and widely used pedagogical frameworks: Bloom's taxonomy and Webb's Depth of Knowledge (DOK). I discuss the potential drawbacks and affordances of the ICAP framework as compared to that of these other pedagogical frameworks. This study adopts the ICAP framework because this relatively new framework has the potential for addressing many problems faced by educators, including course evaluation, student participation, course retention, and student learning outcomes.

Chi and Wylie (2014) framed how students cognitively engage in learning via the ICAP framework. The framework claims that structuring learning around the generation and construction of ideas and knowledge as well as collaborative peer learning improves deeper learning through greater cognitive engagement and deeper inferential learning as compared to instruction that does not require generative and collaborative learning (Chi et al., 2018). Prior research has found that students who take courses designed with higher ICAP levels of engagement perform better on learning measures (Menekse et al., 2013; Wiggins et al., 2017). The ICAP framework translates behaviorist, cognitivist and constructivist learning theories of how humans learn into a hierarchy, which predicts that learning is commensurate to behavioral engagement mode: interactive > constructive > active > passive (Chi, 2009; Chi & Wylie, 2014; Chi et al., 2018). The ICAP framework asserts that the hierarchical modes of overt engagement elicit a differential depth of learning due to the different cognitive processes involved in each mode.

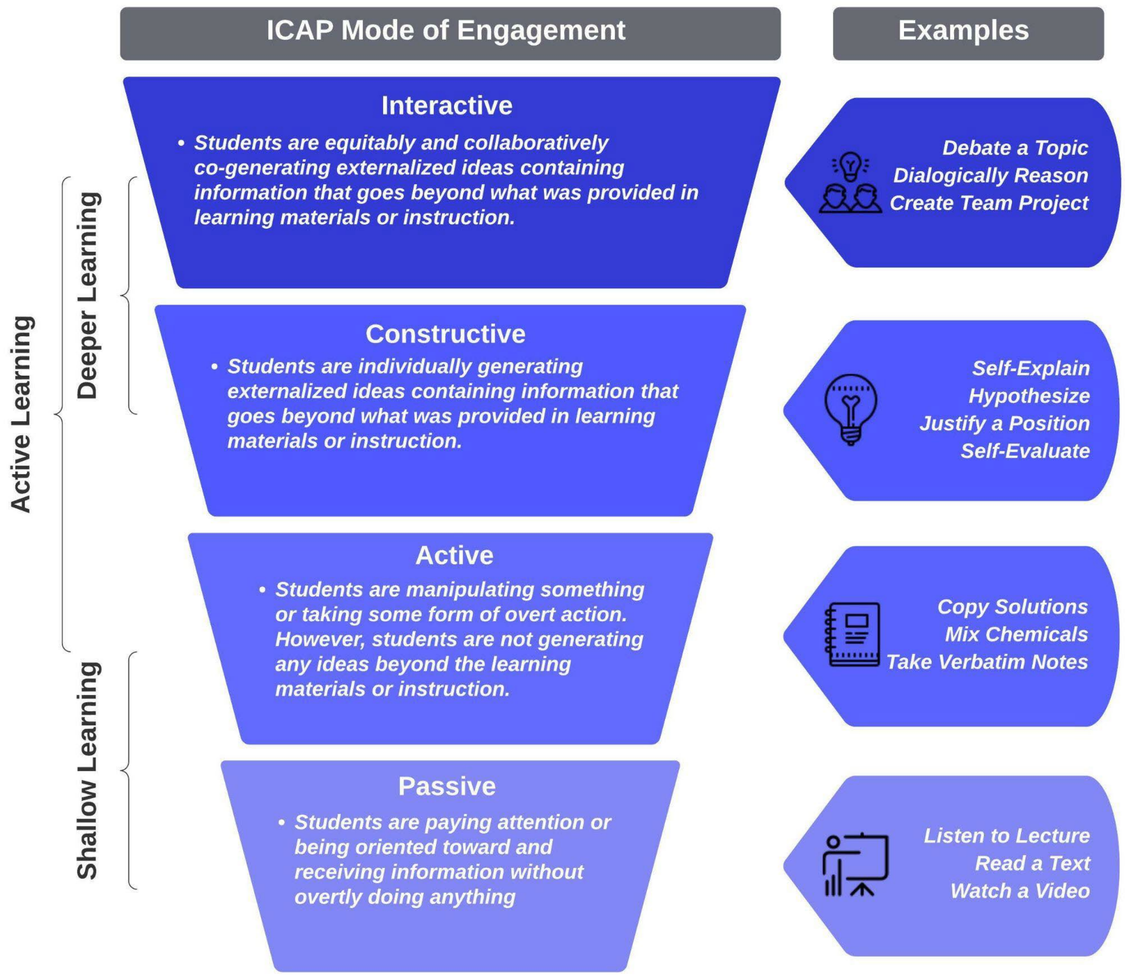
The theory uses overt empirical evidence for determining the ICAP mode engagement. Although students may be covertly engaged in a higher mode of ICAP engagement, instructors, administrators, and researchers are unable to ascertain students' internal cognitive processes. Therefore, the ICAP theory relies on overt manifestations that provide evidence for the cognitive processes that students are likely to be performing.

The main affordance of this framework is that the classification of overt activities has the potential for guiding researchers, instructors, and instructional designers on what kinds of activities and materials would be appropriate for the goals of instruction as well as improve the depth of student learning.

Passive Mode

The passive mode of engagement is defined as paying attention or being oriented toward and receiving information from instructional materials without overtly doing anything else related to learning. For example, a student may be reading an assigned text or listening to a teacher lecture. Students engaged in the passive mode of learning are taking in and storing information, without visibly linking the new information with prior knowledge. The term, passive, refers to the overt behavior students are displaying during the instruction. Students instructed in the passive mode may or may not be thinking deeply about the information since passive learning, by definition, does not include any

Figure 1
The ICAP Framework



observable activity other than paying attention. Although students could be covertly engaging with the learning material in a manner that goes beyond storage, such as elaboration or self-explanation, it is not possible to ascertain such information from only observations of students who are listening to lectures without doing anything else overtly.

It is worth noting that students who are disengaged, as in not paying attention, are not engaged in the passive mode of engagement since they are not attending towards the instruction. Moreover, when students go off-task or engage in discussions that are not instructionally content-relevant, they are also not engaging in any ICAP mode.

Active Mode

The active mode of engagement subsumes the passive mode with the assumption that students are at a minimum paying attention to instruction. The active mode refers to students overtly engaging in activities that activate their knowledge through doing something or overtly manipulating instructional materials as opposed to passively receiving information or instruction (Chi 2009; Chi & Wylie, 2014). The active mode of engagement is operationalized as manipulating something related to learning without constructing any new information or ideas. The cognitive processes hypothesized by Chi (2009) correspond with students activating, encoding, and assimilating new information as well as searching for previously stored knowledge. For example, a student who studies using vocabulary flash cards activates, encodes, and assimilates new information through repeated exposure by oscillating between the word and the definition. In this case, the student is not generating any new information or ideas other than what is presented on the flash card. Covertly, a student studying from a flash card might be generating new ideas or knowledge by relating the definition and the vocabulary term to different prior experiences or using reasoning to parse the word. However, overtly, an observer can only assume that the learner is engaged in the active mode.

The active mode engages students with manipulative activities such as pointing or gesticulating (Alibali & DiRusso, 1999), underlining parts of a text (Katayama et al., 2005), or copying select problem solution steps (VanLehn et al., 2007). Students engaging in the active mode, as opposed to the passive mode, focus more attention on the materials they are manipulating, thereby activating relevant prior knowledge, thus allowing new information to be linked and stored with activated prior knowledge.

It is worth nothing that "active learning" should not be conflated with the active mode within the ICAP framework. "Active learning" refers to a wide array of instructional strategies that are student-centered within the learning sciences literature (Menekse, 2012). "Active learning" includes instructional approaches that may engage students in active, constructive, or interactive modes whereas the active mode precludes collaboration or generation of ideas and knowledge.

Constructive Mode

The constructive mode of engagement subsumes the passive and active modes in the sense that students can only be engaged constructively if they are also overtly paying attention and manipulating instructional materials. The main difference is the addition of generative learning in which students generate and produce externalized ideas containing information that goes beyond what was provided in the learning materials or instruction. This includes creating and inferring new knowledge (at least for the learner and not necessarily universally), integrating new knowledge with old knowledge, re-organizing knowledge, and repairing or accommodating old knowledge.

For example, a student studying a vocabulary word by writing stories that relate the definition of the word to their own life experiences would be overtly engaged in the constructive mode. Other examples of generative activities include self-explanation (Chi et al., 1989; 1994) taking notes in one's own words (Trafton & Trickett, 2001), posing problems (Mestre, 2002), or predicting (Schauble et al., 1995).

Interactive Mode

The interactive mode of engagement subsumes the passive, active, and constructive modes of engagement in the sense that students are required to be overtly paying attention, manipulating instructional materials, and generating externalized ideas that go beyond what is presented in the instruction. The interactive mode adds on the condition of two or more learners undertaking instructional activities and equitably co-constructing and co-generating knowledge and understanding. The interactive mode refers to the interactions between two peers or among a group of students constructing knowledge together through their discourse. To achieve this mode of engagement, the discourse within the group must be primarily constructive, and each partner's contributions engage the group members' contributions, thereby mutually co-generating. In addition, students should be contributing equitably with sufficient turn-taking and shared time on task. Students instructed in the interactive mode go beyond processes such as storing, activating prior knowledge, linking knowledge, or inferring from self, by also inferring from knowledge articulated and constructed by partners and group members.

For example, students learning vocabulary words by equitably co-generating a story or play that uses the vocabulary terms contextually would be overtly engaged in the interactive mode. Examples of interactive activities include debating (Schwarz et al., 2000), asking and answering questions with partners (Webb, 1989), and elaborating on each other's contributions (Hogan et al., 1999).

Ascertaining the Mode of Engagement

Ascertaining the mode of engagement for activities is not always as cut and clear as the previously provided examples. Many instructional activities could be classified as engaging students in multiple modes of engagement. Discriminating among the modes requires careful scrutiny of the ways in which students are cognitively engaged. Here I discuss some examples that may engage students at multiple levels of engagement.

Active vs. Constructive. Students creating concept maps could be constructive or active depending on whether the ideas presented in the map go beyond what is provided in the instructional materials or a lecture. Discriminating between these modes depends on whether students are expected to include ideas and knowledge that go beyond the instructional materials.

Constructive vs. Interactive. Students working together in project-based learning could be either engaged in the constructive or the interactive modes of engagement. If students delegate responsibilities for portions of tasks to only be completed by individual members of a group, the group project would be considered a constructive activity. On the other hand, if students collaboratively and equitably co-construct knowledge on each part of the project, then the group project would be considered an interactive activity.

Another seemingly interactive activity that is actually engaging students in the constructive mode of engagement is jigsaw learning, in which students each engage in discrete tasks and then share individual findings with the group. This activity can also be further downgraded to the active mode of engagement if the discrete portions of the activity do not require students to generate new knowledge.

ICAP Assumptions

The ICAP framework is based on a few assumptions about analyzing learner behavior (Chi & Wylie, 2014). First, instructors may intend for students to engage in a particular mode of engagement, but students may choose to enact unexpected behaviors that may or may not be observable. For example, a teacher may intend for a lecture to engage students in the active mode through note taking, but students may generatively engage in the constructive mode by paraphrasing and summarizing notes on their own accord. Whether students are engaged actively or constructively can be sometimes ascertained by looking at student outputs, in this case student notes. Therefore, at the level of the student, the ICAP framework assumes that the mode of student cognitive engagement is independent of instruction because of this discrepancy between the intended and enacted mode of engagement.

Second, since the previous assumption makes rating instruction more nebulous due to the possible mismatch of intended and enacted, the ICAP framework assumes the rating of greater probability (Chi & Wylie, 2014). The framework assumes that the mode of engagement is more likely based on overt behaviors. Students can be engaging constructively while only listening to a lecture if they are undertaking cognitive processes such as self-explanation or integrating knowledge covertly. However, at the level of the instructor, the framework assumes that the mode of student cognitive engagement is dependent on intended instruction. The theory assumes that based on the design of instruction for lectures, students are more likely to be engaging passively since that is the intended mode of engagement. This assumption can be overruled if the teacher or researcher has access to student outputs that show otherwise. Student level enacted data that differs from what teachers intended takes precedence for ICAP ratings.

Third, the boundaries between the four modes of engagement are not completely discrete. There are some instructional activities that are difficult to classify since they may be either intended for the integration of multiple modes of engagement or dependent on a mixture of cognitive processes. For example, solving problems can be classified as either constructive or active. If the problem asks students to practice a skill such as plugging numbers into an equation via mapping, the problem can be classified as active. If the problem asks students to model a novel contextualized situation and determine how to apply the same equation to the contextualized scenario, this problem may be considered constructive since students have to reconceptualize components of the problem. Another example is a tarsia-based jigsaw puzzle that involves aspects of matching as well as comparing and contrasting. In this case, this task can be rated as 50% active and 50% constructive.

Other Considerations

Assessments. The ICAP framework predicts that students learn more deeply when engaged in higher modes of engagement. If assessments only account for shallow learning, as in questions that only require students to recall information and/or questions that are easy to guess, the ICAP framework may not predict student performance. Menekse et al. (2013) found that students across all modes of engagement performed the same on easier multiple-choice questions but performed commensurate with the levels of ICAP heirarchy on more challenging assessment items.

Domain. The ICAP framework does not predict for certain domains or topics in which students cannot use logical reasoning (deduction, induction, abduction) in order to achieve deeper understanding due to either a lack of relevant schemas or deeper rationale. For example, learning exceptions to grammar rules are difficult to rationalize by learners who do not know them unless the instruction is arranged in a particular manner that allows the learner to infer the patterns of usage.

Related Theoretical Perspectives

Behaviorism

Behaviorism regards learning as the process of operant conditioning in which subjects' observable responses to stimuli condition them to strengthen behavior, which in turn makes a certain response more probable or more frequent, or condition them to weaken behavior, which makes a certain response less probable or frequent. (Skinner, 1965). Learning is mainly accomplished via reinforcement in which subjects are affected by punishments and rewards, whether external or natural, that change behavior. For behaviorists, instruction is essentially reinforcing behavior of learners via the arrangement of stimuli and consequences in the learning environment (Ertmer & Newby, 1993). For example, an instructional sequence stimulus could involve the memorization of vocabulary and the assessment of vocabulary knowledge, leading to a consequential grade on the assessment. Transfer of learning is the generalization of reinforced behaviors that prime learners' responses to identical or similar stimuli. For example, a student who receives a high score on the vocabulary test may subsequently be more motivated to study more diligently on vocabulary terms. At the same time, according to Schunk and Meece (1992), behaviorist principles cannot account for acquisition of higher-level skill sets or deeper learning, such as those involved in problem solving and inference generation. A major application of behaviorism in learning is in producing observable and measurable outcomes for students, emphasizing the practice of learning, sequencing instruction, and using reinforcement to impact student performance via feedback (Ertmer & Newby, 1993).

Behaviorism and Objectivist Epistemology. Behaviorism is rooted in objectivist epistemology, which posits that the world is real and external to humans as well as independent of human experience (Jonassen, 1991). Objectivists regard learning as the conceptualization and categorization of abstract symbols that correspond to the one true and correct reality (Vrasidas, 2000). To objectivists, there is only one valid understanding. Knowledge and meaning exists objectively and independently and is external to the knower (Cronje, 2006). Since there is only one reality, the structure of the world can be modeled for learners so that the mind can mirror the one reality via learning by mapping entities and concepts (Jonassen, 1991). Therefore, instruction is a transfer of objective knowledge from an instructor to the learner (Vrasidas, 2000). The role of education is for students to learn about the real world, and the role of the teacher is to instruct events for students, who are discouraged from making their own interpretations (Jonassen, 1991). In relation to behaviorism, since knowledge is objective, optimal instruction should affect learner behavior in a manner that maximizes their learning of these objective truths, thus reinforcement.

ICAP and Behaviorism. The ICAP framework relates to behaviorism in terms of the overt behavior of students. In the passive mode, students are attending to the instruction directed towards them whether direct instruction or reading as opposed to being disengaged or not paying attention. In the active mode, students are attending to instruction as well as physically manipulating something related to the instruction such as measuring an Erlenmeyer flask or writing down what the teacher is saying. In addition, the constructive and interactive modes assume students are attending to instruction. These behaviors are directly measurable and are the basis of categorizing instruction within these modes of engagement.

In terms of student learning, the application of behaviorism to the ICAP framework is somewhat limited since the framework does not concern itself with external reinforcement but rather on how students are cognitively engaged in learning. Skinner (1968) addressed how behaviorist principles relate to instructional programming, motivation, discipline, and creativity but did not directly relate behaviorist principles to student cognitive engagement since cognitive processes are not observable. At the same time, behaviorism does relate to some aspects and types of activities that engage students in varying ICAP modes in terms of learning transfer (e.g., practicing test questions, conducting laboratory experiments, writing tasks, etc.). These forms of tasks are likely to generalize and affect student performance on assessments (e.g., examinations, essays, and lab examinations) that are directly related to or similar to these aforementioned activities. Students who form habits from working on generalizable tasks, are likely to perform well on related assessments and measurements. For example, students engaged in the passive and active modes may perform proportionally better on assessments that do not test for inferential learning since students can generalize their learning to performance on non-inferential questions efficiently.

One major aspect the ICAP theory inherits from behaviorism is the focus and reliance on observable actions of students as a foundation for assessing and evaluating the effectiveness of learning. Skinner (1968) asserted that though there are cognitive processes that occur internally, only externally observable behaviors can be objectively measured. One major difference between the ICAP theory and behaviorism is that ICAP affords the measurement of student cognition in terms of how they are likely to be cognitively processing information based on the behaviors that are seen and expected of students. For example, within a classroom, a student who is asked to analogize could either be generating new information or repeating an analogy that is already known from prior knowledge. But as a whole, the class is more likely to be generating new information.

Cognitivism

Cognitivism regards learning as the processing, storing, and organizing of information in learner memory (Ertmer & Newby, 1993). Cognitivists are interested in what learners know and the structure of their knowledge stored inside of their brains. They emphasize how learners process information and incorporate it into their pre-existing database of knowledge. Cognitivists acknowledge the role of how teachers create environments for learning much like that of behaviorism; however, cognitive theories go beyond instruction and focus on how students cognitively process the instruction they receive (Schunk, 2012). Students learn by selecting and attending to portions of learning stimuli they encounter, transforming and rehearsing information in relation to prior knowledge, and organizing knowledge in a meaningful way (Mayer, 1996).

Cognitivism is rooted in information processing theories of how learners encode information into memory and retrieve it as needed (Shuell, 1986). Cognition is the activity of information processing mechanisms (Neisser, 2014). Although not all researchers accept this model, the two-store (dual) memory model is the most commonly used information processing theory, which was influenced by the emergence of computers in the 1950s (Friedman et al., 2013). In this dual memory model, students attend to stimuli and transfer them into short-term or working memory (WM; Terry, 2015), which then is rehearsed or related to long-term memory (LTM). WM is limited in duration and capacity; most people can only hold a small amount of information (i.e., about seven items) in their WM (Miller, 1956). This capacity can be increased via chunking, or the combination of information in a meaningful fashion, e.g., a phone number, though seven digits can be chunked into two larger numbers. Information that is processed in the WM is then encoded and stored in the LTM as episodic or semantic memory (Tulving, 1992; 1993). Episodic memory includes information associated with particular memories from a person's life that is highly personal, whereas semantic memory involves general information and knowledge that is not tied to a particular personal context, such as concepts learned in school.

Encoding is the process of preparing information in the WM for storage in the LTM via making information meaningful and relating it to prior knowledge already stored in the LTM (Schunk, 2012). There are three major influences on the effectiveness of encoding, including organization, elaboration, and schemas.

The Gestalt theory showed that the organization of material could enhance encoding by grouping information into chunks (Katona, 1940; Miller, 1956). Some examples of organization include hierarchies and mnemonics. Elaboration is the process of expanding upon information in the WM by adding to or linking to information already stored in the LTM (Schunk, 2012). For example, a student learning about the law of cosines, "$c^2 = a^2 + b^2 - 2*a*b*\cos(C)$," can link and relate the information to the already known Pythagorean theorem, "$a^2 + b^2 = c^2$." Schemas are meaningful structures that organize large quantities of information into a system (Schunk, 2012). They assist in the completion of routine sequential actions that may be different in action but, in essence, similar in form (Cooper & Shallice, 2006). Examples of schemas include processes for solving types of mathematics problems and laboratory procedures in a chemistry experiment.

Learners store encoded and organized information from the WM into the LTM as declarative or procedural knowledge. Declarative knowledge, which includes facts, beliefs, opinions, theories, ideas, and attitudes about oneself (Gupta & Cohen, 2002), is acquired by learners whenever propositions within a propositional network are stored. Propositions are the smallest unit of information that people judge to be true or false and are the basic units of knowledge stored in LTM (Anderson, 2005; Norman & Rumelhart, 1975). The ACT-R theory (Anderson, 2005) proposed that propositions are formed through the combination of two nodes: subject and predicate nodes. For example, the subject of puzzles is usually linked to the predicate of solving to form a proposition of solving puzzles. This same proposition can be linked into a more complex proposition through the linking of another proposition, e.g., solving puzzles takes time. Learners create propositions whenever they receive new information and process them in WM. These propositions are then stored in LTM via spreading activation (Anderson, 1983), in which propositional nodes in the WM associated with prior knowledge nodes in the STM are retrieved and activated. For example, if learners are asked to describe something that takes a long time, they may be primed to respond with the proposition, solving puzzles. Procedural knowledge is stored in a similar fashion to declarative knowledge, but the focus of such knowledge is on the sequence or rules of a procedure (Anderson, 1982).

Once LTM is stored, learners retrieve knowledge in order to perform actions that are informed by prior knowledge. Whenever information is retrieved, the ideas that are in question to be retrieved enter the WM as propositions; and associated information in LTM is accessed via spreading activation (Anderson, 2005). In essence, retrieving knowledge from LTM is rehearsal and would strengthen the future activation of propositional nodes. Organization and elaboration both enhance the effective processing and retrieval of LTM (Schunk, 2012). Moreover, meaningfulness is a key factor in LTM storage. Though repeated rehearsal may store information in the LTM, the strength of propositional networks depends on how they are connected to other propositional networks, a.k.a. meaningfulness. For example, students can repeatedly rehearse the Pythagorean theorem as "$a^2 + b^2 = c^2$" to store it in LTM but can enhance the network by elaborating the meaningful fact that the theorem is useful for finding a missing side of a right triangle. The addition of meaning now connects the propositional network of the procedural Pythagorean theorem with the declarative networks of triangles and missing values. Therefore, elaboration increases the likelihood that information is stored in LTM by associating propositional networks (Stein et al., 1984). Organization improves LTM by expanding a propositional network by breaking information into parts and creating relationships among those parts (Schunk, 2012). For example, the organization of trigonometric rules with chunked mnemonics such as "Soh Cah Toa" can help learners to store the relationships of sine (S), cosine (C), and tangent (T) to the opposite (o), adjacent (a), or hypotenuse (h) sides of a triangle.

Learning is considered effective when students are able to effectively transfer and apply prior knowledge, learning, and skills to new learning (Schunk, 2012). Transfer is dependent on the strength, complexity, and network connections of proposition nodes. Based on the cognitive model, the goal of instructional design is to create efficient learning that allows learners to activate relevant prior knowledge in LTM and activate relevant and meaningful propositional networks.

Two techniques that are used to achieve efficiency of knowledge transfer are simplification and standardization (Ertmer & Newby, 1993). Effective instruction simplifies knowledge into basic building blocks that are chunked in ways, so that proposition network activation is efficient. One strategy is the use of advanced organizers that direct learners' attention to important concepts, highlight relationships among ideas, and link new material to what students know (Faw & Walker, 1976). This strategy is based on Ausubel's (1977; 1978) theory of meaningful reception learning that hypothesized that learning is meaningful insofar as new information bears a systematic relationship to concepts stored in LTM. Meaningfulness elaborates, modifies, and expands information in the proposition network.

Information in instruction should be presented in ways that allow students to relate new information with known information and the understanding of the uses of such information (Schunk, 2012). Learning should be structured to build on prior knowledge stored in the LTM. Moreover, feedback, in terms of knowledge of results, supports learning through the rehearsal and organization of propositional networks (Ertmer & Newby, 1993). Learners benefit from structuring, organizing, and sequencing information for facilitating optimal chunking of propositions in the WM. In addition to feedback, learning environments that allow and encourage students to make meaningful connections with prior LTM expands propositional networks via elaboration. There are a number of cognitive strategies in line with cognitivist theories, such as the use of analogies, comparing and contrasting, outlining, concept mapping, summarizing, self-testing, self-explaining, etc.

Cognitivism and Rationalist Epistemology. Cognitivism is based on rationalist epistemology. According to rationalism, knowledge is innate, and humans derive their conception of knowledge purely from reason and intuition (Schunk, 2012). Knowledge arises out of experience but is not grounded in it. Instead, meaning is revealed to people via their senses, and learning is a process of discovering ideas by reflecting via logical reasoning (Schunk, 2012). In relation to cognitivism, information gathered from the senses are processed via logical reasoning and then subsequently stored.

ICAP and Cognitivism. The higher modes of the ICAP framework are heavily influenced by information processing theories and rationalist epistemology. In essence, generative learning, or the interactive and constructive modes of engagement, assumes the involvement of information processing theory with students engaging in learning using generative cognitive processes of encoding such as elaboration, organization, and schema generation. Only when instruction shows evidence of students engaging in these encoding processes, the activity can then be considered as generative learning. In addition, the cognitive focus on logical reasoning is an assumption in the ICAP framework that is the basis for why instruction that does not follow natural logic, such as exceptions in language rules, does not predict student learning.

Additionally, though the ICAP framework inherits concepts of cognition and reasoning from rationalist epistemology, the higher modes of the ICAP theory disagree with rationalism in terms of the nature of knowledge. The rationalist approach views knowledge as innate; in contrast, the higher modes of the ICAP framework adopt the constructivist epistemological approach, which views knowledge as constructed.

Constructivism

Constructivism regards learning as the process of interpreting information and constructing new meaning in relation to prior knowledge, an individual meaning-making. According to constructivism, though an external reality exists, the relevant reality is contingent on the mind of the knower, who subjectively interprets and constructs his or her version of reality via experiences and interactions with the external world (Crotty, 2020). Constructivists contend that learners cannot know the exact truth or knowledge but rather interpretations of them.

Learning is the process of interpreting phenomena and the construction of mental models that represent the phenomena (Driscoll, 2000; Fosnot, 2005). Knowledge and meaning are a created in an interpretive process that is dependent on the experience and perspective of the knower (Cronje, 2006). Learning is therefore very personal, since the individual is the constructor of meaning. Moreover, knowledge is created by interpreting and connecting it to existing knowledge and context (Clark, 2018). Student learning takes on a dual role - public and private - in which meaning and knowledge are constructed when learners interact with the physical world and make sense of that interaction internally (Henriques, 1997).

Though constructivism has roots in Kantian philosophy, much of the application of constructivism to learning theories is based on the works of John Dewey. Dewey asserted that the best learning occurs: a) when students learn by experience and have the opportunity to partake in their learning; b) when students are involved in social and interactive processes; c) when students are part of a social group that influences their own learning; and d) when students are challenged to use creativity for arriving at individual solutions for difficult problems (Dewey, 1933, 1986). Dewey (1923) argued that knowledge is constructed through active experience and processing. Accordingly, Dewey's ideas are often cited as a major source for "active learning" methods, which applies to the active, constructive, and interactive modes of ICAP engagement, see figure 1.

Constructivism and Empiricist Epistemology. Constructivism is rooted in empiricism (Matthews, 1992), in which experience and perception are the only source of knowledge (Schunk, 2012). This post-positivist epistemology contends that learners cannot know the exact truth or knowledge but rather interpretations of them. In contrast to objectivism and rationalism, students construct their own knowledge (Piaget, 1952). Meaning is not created, as in subjectivism, but rather constructed since the objective world exists externally to constructivists. Learners construct their own interpretations of the nature of the objective world. Knowledge is not accumulated or attained, but rather the result of active cognizing by individuals (Von Glaserfeld, 1984). The shift from objectivist to constructionist epistemologies shift the focus of learning from teachers to the students.

ICAP and Constructivism. The main implication of constructivism and empiricist epistemology that applies to the ICAP framework is that students should be actively involved in constructing their own learning. This applies for the active, constructive, and interactive modes of engagement. In a sense, constructivists are concerned mainly about whether students are actively learning and constructing meaning.

Dewey's constructivism relates to "active learning" strategies, which are very diverse including widely differing cognitive tasks such as matching (Menekse et al., 2013), self-explaining (Chi et al., 1994), generating predictions (Klahr & Nigam, 2004), copying (VanLehn et al., 2007), or underlining and highlighting (Igo et al, 2005). However, "active learning" does not discriminate differences among the three higher modes of cognitive engagement. At the same time, there are two branches of constructivism that do address the differences among the ICAP modes of engagement within "active learning".

Branches of Constructivism

There are two constructivist branches that differentiate how students cognitively engage with different "active learning" strategies, including cognitive constructivism and social constructivism. In the following sections, I further discuss cognitive and social constructivism and their relationships with and implications for the ICAP framework.

Cognitive Constructivism. Cognitive constructivism is a theoretical perspective coined by Jean Piaget that integrates information processing theories with constructivism. Piaget theorized that human cognitive development has four stages: a) sensorimotor in which children construct their world primarily through their own senses, physical activity, and basic language skills; b) preoperational in which students begin to grasp symbolic function and develop more complex language skills while not being able to effectively take on the perspective of others; c) concrete operational in which children begin to use logical reasoning; and d) formal operational in which teenagers to adults further develop logical reasoning with the addition of abstract and inferential reasoning (Powell & Kalina, 2009). According to Piaget, learning happens through assimilation and accommodation in which learners find a way to integrate and balance new and old information, also known as equilibration (Piaget, 1952; Wadsworth, 2004). Learning happens in two ways: a) when the learner encounters new information and readily assimilates it into their existing mental model, a.k.a. assimilation; and b) when the learner encounters cognitive conflict in trying to make sense of new information and accommodates for the new information in relation to the old by reorganizing the existing mental model, a.k.a. accommodation (Powell & Kalina, 2009).

In terms of information processing, learners do not transfer knowledge from the external world but rather build personal interpretations of the world based on individual experience and interaction (Ertmer, & Newby, 1993). Learners are active agents who must discover the principles for themselves (Geary, 1995). Memory is no longer for the storage of information and facts but rather an accumulative history of interactions and interpretative constructions of knowledge and information (Ertmer & Newby, 1993). Instead of retrieving intact knowledge, learners assemble meaning through the interaction of prior constructed knowledge in the LTM with the situation at hand. The cognitive constructivist perspective is a shift away from learning as the acquisition and recall of fixed concepts or details towards the flexible application and integration of knowledge. Accordingly, learners need rich learning environments that allow for the active construction of knowledge (Schunk, 2012).

ICAP and Cognitive Constructivism. Cognitive constructivist ideas of how individuals learn align very closely with the constructive category of the ICAP framework in the sense that knowledge is constructed by individual learners via cognitive processes. In addition, cognitive constructivism also relates strongly with the interactive mode with only major differences in terms of the role of society in the construction of knowledge and meaning.

One major consideration from Piaget's theory is that the ICAP framework is probably more applicable in the context of students who have reached formal operations, approximately 12 years old and up. Since the ICAP framework predicts improved deeper inferential learning, students may begin reaping increased benefits from instruction at the higher modes of engagement starting in middle school.

Social Constructivism. Social constructivism is a theoretical perspective coined by Lev Vygotsky that moves the focus on individual cognitive construction towards social construction of knowledge and meaning. Social constructivism extends the principles of constructivist learning theory into the context of interactive learning environments in the sense that meaning making is not individualistic but rather influenced by the social origin of meaning (Crotty, 2020). Therefore, all constructed interpretations of the external world by individuals are influenced by and constructed within society. According to Vygotsky's theory, social interactions are critical because knowledge is co-constructed and negotiated among groups through the transmission of cultural tools such as language and symbols (Meece, 2002; Vygotsky, 1978). Vygotsky's research stressed the importance of social context in the learning process through major theories such as the zone of proximal development (ZPD), scaffolding, and cooperative learning (Vygotsky, 1962). The ZPD refers to the zone where effective learning occurs when a learner is involved in learning that is a little too difficult to perform and receives critical help and support from others in order to reach the next level of understanding (Powell & Kalina, 2009). Teacher, expert, or tutor support for students within the ZPD in this assisted learning process is called scaffolding, in which students are given external support for accomplishing tasks and learning that would be otherwise too difficult. To Vygotsky, social interaction plays a pivotal role in the process of cognitive development and learning. In addition to learning from teachers, learners also learn by interacting with one another. Meaningful interactions include negotiating socially, debating points, adding upon evolving ideas, and offering alternative explanations and perspectives with one another while tackling authentic tasks (Lave & Wenger, 1991; Vrasidas & McIsaac, 1999; Vygotsky, 1978). According to Vygotsky, cooperative learning via interactions is essential for creating deeper understanding through the use of language (Powell & Kalina, 2009).

ICAP and Social Constructivism. The ICAP framework's interactive mode of engagement aligns very closely with social constructivist ideas of how individuals learn within society. The interactive mode assumes that the deepest learning occurs when learners engage in generative cognitive processes in tandem with one another. The ZPD also points to the co-generation of new ideas and co-construction of knowledge since learning happens at the locus where students are challenged to learn beyond what they already know.

Related Frameworks

Bloom's Taxonomy

A major theoretical framework that is often juxtaposed with the ICAP framework is Bloom's Taxonomy (BT: Bloom et al., 1956; Anderson & Krathwohl, 2001). Bloom's taxonomy is a hierarchy of cognitive processes represented cumulatively in which each higher level assumes mastery of each lower level (Krathwohl, 2002). Bloom's revised classification has a hierarchical assumption: Knowledge < Comprehension < Application < Analysis < Evaluation < Synthesis. According to Bloom, learners need to acquire and remember knowledge in order to subsequently comprehend/understand, apply, analyze, evaluate, and create/synthesize knowledge. Bloom's taxonomy was conceived as a means for creating standardized test items for measuring educational objectives (Krathwohl, 2002). In essence, the taxonomy takes a cognitive approach to organizing and categorizing what students learn. As such, Bloom's taxonomy has been applied by educators and instructional designers in the development of instructional materials and assessments that align with the achievement of all cognitive levels of the BT hierarchy (Anderson & Krathwohl, 2001).

In a sense, Bloom's taxonomy is similar to ICAP since cognitive tasks in application, synthesis, and evaluation tend to elicit the Constructive mode more than active or passive modes. But a major distinction between the two is that Bloom's taxonomy applies to planning and preparation (the back end), whereas the ICAP framework applies to planning, preparation, instruction, and evaluation (the front end and back end). A teacher may use Bloom's taxonomy to select what kinds of items should be in a test, whereas a teacher may use the ICAP framework for test creation and for deciding how students should cognitively engage with the curriculum. An evaluator can use Bloom's taxonomy to rate an assessment but cannot use the taxonomy to effectively rate how instructor's choice of instructional activities will cognitively engage students. On the other hand, the ICAP framework can be used to assess an instructor's pedagogy on all levels of instruction. Bloom's taxonomy is limited to the creation and evaluation of assessments and commensurate learning materials. The ICAP framework is also applicable to how instruction is structured and performed in manners that allow for the construction and generation of knowledge.

At the ground level, there are only a few major differences in terms of instructional approach. Instruction that aligns with higher levels of Bloom's taxonomy is likely to achieve higher levels of the ICAP mode but not certainly. What the ICAP framework adds is how learners are cognitively engaging in terms of generative activity. Students can apply, evaluate, and synthesize in ways that may be rehearsing knowledge that is already known to the learner, as in no generative learning; though these actions are at the higher end of Bloom's hierarchy, they would fall in the lower portion of the ICAP hierarchy. The higher modes of the ICAP framework necessitate learners interpreting and constructing/generating different ideas for how to apply, evaluate, and synthesize. Further, Bloom's Taxonomy does not incorporate social constructivist ideas on the value of interactive learning for deeper learning.

Webb's Depth of Knowledge

Webb's Depth of Knowledge (DOK) is a more recent taxonomy that categorizes cognitive complexity (Hess et al., 2009). Webb (1997, 1999) sought to reframe how content and assessment can be aligned based on the complexity of content (Webb, 1997; 1999). There are four levels of increasing order of complexity: a) level 1, recall and reproduction; b) level 2, application of skills and concepts; c) level 3, strategic thinking; and d) level 4, extended thinking. At level 1, students are expected to engage in recalling facts and concepts as well as perform routine procedures via reproduction e.g., memorizing vocabulary or copying down class notes. At level 2, students are expected to apply skills and concepts by using information and conceptual knowledge for selecting procedures for tasks and problems that require more than two steps or decision points (e.g., solving similar problems to one presented in class). At level 3, students are expected to reason or develop plans for approaching complex problems, that usually have more than one possible answer by involving inferential reasoning (e.g., solving complex problems such as Olympiad mathematics questions or abstract physics problems that require modeling of the problem). At level 4, students are expected to perform multi-disciplinary investigations or apply skills and concepts in real world problems that require research, problem solving, and the consideration of multiple conditions of the problem or task (e.g., a well-designed interdisciplinary project-based learning experience). The major difference between Bloom's taxonomy and Webb's DOK is that BT is focused on the taxonomy of cognitive processes involved in learning whereas Webb's DOK is focused on the depth and complexity of learning experiences (Hess et al., 2009).

The DOK relates to the ICAP framework in the sense that the DOK is concerned with the depth of cognitive processing and whether students are likely to learn more deeply. The higher levels of DOK certainly involve generative learning and level 2 may or may not involve generative learning depending on how students approach the application of skills and concepts. Like BT, the DOK can be used for creating differential instructional materials and assessments. However, unlike BT, the DOK can be used for evaluating instructor pedagogical practices in instruction since the framework is concerned with the depth of cognitive processing of students, albeit in terms of the complexity of tasks, not generative collaborative learning.

The ICAP framework differs from the DOK at a few levels. First, generative learning discretely defines the levels of the ICAP framework. In contrast, for Webb's DOK, level 2 may or may not engage students generatively whereas the ICAP distinguishes the lower and higher levels based on whether students are engaged in generative learning. While the DOK levels 3 and 4 probably presume generative learning, this is not explicitly expected of students. Second, the ICAP framework directly applies implications of cognitive and social constructivism in the higher modes whereas the DOK does not necessarily assume that students are constructing knowledge individually or collaboratively. Project-based learning often includes collaboration among learners; however, the 4th level of DOK may fall short of the collaborative assumptions of the interactive mode such as co-generation, co-construction, and equitable contribution in collaboration.

Some considerations for further development of the ICAP framework include the differentiation of DOK levels three and four. These two levels both possibly involve generative learning; however, level four assumes a high level of cognitive complexity for deeper learning. Perhaps future research on the ICAP framework may benefit from exploring differential levels of task complexity within each mode of the ICAP framework.

CHAPTER 3

Background Literature

Online Learning

The affordances of technology have driven the creation and evolution of different formats of online learning that address the need for accessible, high-quality education. Three formats for online learning have emerged, including asynchronous online instruction (AOI), synchronous interactive online instruction (SIOI), and blended synchronous learning (BSL). In this section, I discuss differences among these formats of online learning and how they relate to student outcomes and perceptions.

AOI refers to the delay in communication among students and instructors in a course (Skylar, 2009). This communication is typically relayed using an LMS, such as Blackboard, Moodle, or Canvas (Kruger et al., 2015). The major feature and advantage of AOI is that the format of instruction affords students the flexibility to complete coursework anytime and anywhere (Ahmad & Bokhari, 2013).

SIOI refers to the immediacy of communication among students and instructors in a course (Skylar, 2009). SIOI typically uses a web conferencing application, such as Zoom or GoToMeeting, to facilitate synchronous online instruction (Al-Samarraie, 2019). The major feature and advantage of SIOI is real-time communication and simultaneous viewing of learning materials (Bower, 2011; Skylar, 2009). Studies have found that participants learning from SIOI as compared to AOI tend to participate more and experience higher task/course completion rates (Chen & You, 2007; Hrastinski, 2010; Mabrito, 2006).

In a mixed-methods study, Ward et al. (2010) found some key findings regarding student and faculty perceptions of AOI and SIOI. They found that students rated SIOI significantly higher than AOI ($p < .001$) in dimensions of instructional effectiveness, including quality and amount of content learned, cooperation among students, “active learning”, and time on task. In addition, Ward et al. found that instructors generally had favorable views of SIOI. On the other hand, students rated AOI ($p < .001$) higher than SIOI for ease of access. This research suggests that both formats of online learning may provide exclusive benefits for learners.

During the COVID-19 pandemic, many universities, including the university within this study, shifted their F2F classes into BSL, which combines aspects of AOI and SIOI (Alexander et al., 2019; Johnson et al., 2016), potentially providing students the accessibility of AOI and the instructional effectiveness of SIOI. In this format of instruction, students and instructors communicate in varying formats, including video conferencing and discussion posting, in addition to static websites for communication (Bower, 2011). Yamagata-Lynch (2014) found that students learning with BSL experienced higher levels of participation while maintaining a flexible learning environment as compared to that of AOI. In addition, BSL can establish an environment with rich teaching, social, and cognitive presence that can positively impact participation level (Szeto, 2015).

In a case study that compared student perception differences among F2F learning, AOI, and BSL, Bower et al. (2015) found that approximately three-quarters of students preferred BSL over AOI and F2F learning. In addition, Bower et al. found that students reported the most “active learning” (e.g., question and answer sessions, role plays, collaborative evaluation, whiteboard exercises) in BSL. Bower et al.'s findings on "active learning" may suggest improved learning with BSL compared to that of AOI. However, the quality of "active learning" needs more careful evaluation since "active learning" does not differentiate among the higher levels of ICAP engagement (i.e., active, constructive, or interactive modes), which in turn may have differential effects on learning outcomes.

The literature shows that there are clear differences in course format that may affect student learning, course retention, participation level, and course outcomes. Taking these considerations into account, this study examined the differences among formats of instruction, AOI, F2F, and BSL, that may influence student outcomes. I did not examine SIOI since few postsecondary online courses fall strictly within this category. Furthermore, the data in this study did not include any SIOI classes.

LMS and Course Participation

LMS use is now ubiquitous in postsecondary institutions and affords instructors and administrators the ability to monitor student learning participation and progress (You, 2015), discover meaningful patterns (Gašević et al., 2015), and adjust instructional strategies accordingly (Dietz-Uhler & Hurn, 2013). In exploring the impact of LMS on education institutions, Reyes (2015) states that LMS data could be used for better understanding what factors may affect course retention and for informing changes instructors could make to course design and instructional strategies. Accordingly, researchers have begun to use LMS data for examining online learning success. Studies have focused on measures of student participation such as the number of content views, frequency of logins, and time spent reading pages to explain individual learning outcomes (Henrie et al. 2018; Morris et al., 2005; Qu & Johnson, 2005). Moreover, researchers have reported that significant differences in student performance are related to participation level in online courses and their respective LMS (Johnson, 2005; Morris et al., 2005; Wang & Newlin, 2002; You, 2016). Considering student participation level, as measured by LMS use, may have an impact on student learning, this study also focused on the impact course design may have on student LMS participation, or RQ 1.

Course Retention

Dietz-Uhler and Hurn (2013) note that there is a lack of a common definition of course retention. They found that definitions range from liberal (percentage of students who do not withdraw from a course or students who pass the course) to conservative (percentage of students starting the first module or assignment and remaining in the course from start to finish). This study adopted the latter definition of retention since the focus is on evaluating course design and not other factors (e.g., personal reasons) for course withdrawal.

Research on online course retention has found that one of the main factors for student withdrawal from a course or program is how courses are designed (Bornschlegl & Cashman, 2019). Pierrakeas et al. (2004) found that poor teaching methods and learning materials related to students' decision to withdraw from a course. In addition, Meyer et al. (2009) found that good course design influenced students' motivation to stay enrolled in AOI courses. Accordingly, this study also examined the relationship between the evaluation of course design and course retention, or RQ 4.

Design and Evaluation of College Courses

Researchers have found that universities and colleges use rubrics to evaluate course quality (Jaggars & Xu, 2016; Yuan & Recker, 2019). Prior research has shown that course evaluation with rubrics has the potential for predicting student course retention. Of the available options, the Quality Matters higher education (QMHE) rubric (Quality Matters, 2018) is one of the most widely used rubrics used to evaluate courses in higher education (Jaggars & Xu, 2016). The QMHE rubric consists of eight standards (e.g., assessments and measurement, instructional materials, learning activities, and learner interaction) that detail what should be included and how the course design should be structured.

In a non-experimental study that explored the application of an early version of the QMHE rubric to online courses in an undergraduate psychology and statistics course, Dietz-Uhler and Hurn (2013) found that courses that met the evaluation standards of the QMHE rubric resulted in high course retention rates, 95.5% over 11 course offerings. Although Dietz-Uhler and Hurn do not provide the empirical causal evidence for evaluating the effect online course design may have on student course retention, the study points to the need for more rigorous research on this relationship.

A recent review of course design evaluation instruments compared and contrasted six widely used course design rubrics, including the QMHE rubric, in universities and found that the evaluation of course design quality focused mainly on course structure, communication, and assessment (Baldwin et al., 2018). While most of the rubrics encourage collaboration and peer to peer interaction, none evaluate the quality of cognitive engagement of course activities and instructional materials. Some of these rubrics encourage "active learning" techniques as a generality but without specifics on the differing quality of activities for cognitively engaging students in learning, which, in contrast, the ICAP framework affords. The Quality Matters higher education (QMHE) rubric standards 5.1 and 5.2 state that courses should be aligned to promote student "active learning" but the QMHE rubric gives no guidance on what kinds of "active learning" techniques and activities would better promote student learning (Quality Matters, 2018). There is a clear need for a theory-based approach for designing online courses and for evaluating the quality of course design for improved student learning. The ICAP instructional rubric offers a possible solution to fill this gap in course design and evaluation.

CHAPTER 4

Methodology

Research Questions & Rationale

1) To what extent are learning activities and learning materials, as measured by the ICAP instructional rubric, associated with student participation in course LMS?

2) To what extent do learning activities and learning materials, as measured by the ICAP instructional rubric, predict student final exam scores?

3) To what extent do learning activities and learning materials, as measured by the ICAP instructional rubric, predict student course grades?

4) To what extent are learning activities and learning materials, as measured by the ICAP instructional rubric, predictive of student course retention?

5) To what extent are learning activities and learning materials, as measured by the ICAP instructional rubric, predictive of student course attrition?

The ICAP framework predicts student deeper learning (Chi & Wylie, 2014), which is only directly related to research question 2, predicting final exam scores. Nevertheless, I chose to also study the other four relationships, (i.e., course participation, course grades, course retention, and course attrition) since these measures of student outcomes are of great interest for the field of education data analytics and education administrators. Moreover, the ICAP framework, specifically, has not been previously researched in this manner whereas studies have been conducted on the relationship among these four student variables and “active learning”. These results potentially contribute new findings for the fields of learning sciences, learning analytics, and higher education administration.

Research Design

The design of this study is based on the application of multivariate regression analysis to secondary data. The participants of the study include all those who self-registered into F2F, AOI, and BSL sections of an introductory chemistry course at a large public university in the southwest United States.

Data

This dissertation study examines data collected from Fall of 2019 until Spring of 2021. Instructional and student data were sourced from a single introductory chemistry course that were routinely collected and stored within the university's learning management system software, Canvas, and the university's human resources software, PeopleSoft. Considering student data was collected via database and not sourced from participants, consent was not required from participants. All student personally identifiable information from PeopleSoft as well as Canvas was de-identified and coded into alphanumeric values. Fifteen instructors of the course were identified through the introductory chemistry course listing within the time frame and recruited by email; consent was received for accessing course data. The data included 3277 student participants in 13 course sections with 5 instructors. In terms of instructional mode, 3 course sections were F2F, 4 course sections were AOI, and 6 course sections were BSL. In large part, for each instructor, course shells did not vary from semester to semester. There were differences in terms of timing and announcements from semester to semester; however, assignments were consistent. One instructor taught two variations of the course that varied by mode, AOI and BSL. Therefore, of the 13 course sections taught by 5 instructors, the data included 6 variations of the introductory course.

Student data from PeopleSoft included demographic information and measures of student prior academic performance that are further described in the next section. Instructor data from Canvas included each instructor's course shell for their course sections. Overall, the Canvas data included syllabi, course schedules, video and written resources, announcements, and assignments. Types of assignments within the Canvas included homework and discussion posts. Types of assignments also graded within the course but were not consistently found in Canvas course shells included recitation and laboratory (lab). Recitation is an instructional activity designed for students to review concepts from lectures by collaboratively solving skill and concept problems. Laboratory included a few components: pre-lab quizzes, experiments, post-lab assignments, inquiry investigations, investigation lab reports, lab final exams, and a lab research poster project.

Instrumentation

The ICAP Instructional Rubric (Appendix A) is an instrument for measuring the ICAP dose of instruction. Adapted from the work of Stump et al. (2017) that rated instructional videos using a coding procedure based on the ICAP framework, an ICAP dose refers to the number of minutes students spent engaged in a particular ICAP mode. The rubric builds upon previous work by creating a specific procedure that guides raters for assigning ratings and calculating dose. Ratings should be based on empirical data such as: course syllabi, course schedules, LMS course shells, classroom observations, recorded videos, lesson plans, lesson materials, student work products, instructor interviews, and student interviews.

ICAP dose should be calculated based on the finest grain of data available. For example, a teacher's lesson plan might enumerate all the intended activities, materials, time spent; however, the actual lesson may have been enacted with deviations from the lesson plan. Students' work products may show that the teacher's intended ICAP mode of engagement was not achieved. In this case, the ICAP dose should be calculated based on how the lesson was enacted and not on the teacher's lesson plan. For cases in which enacted data is lacking, the ICAP dose should be calculated based on the intended lesson. An example of coarser to finer grained data in a college course is as follows, respectively: syllabus, course schedule, course LMS layout, lesson plans, instructional materials, observations of instruction, teacher interviews, student interviews, audio recordings of classes, video recording of classes, and student work products. Ideally, rating instruction with the rubric includes influence from all levels of data available to elucidate ICAP dose ratings that best represent the instruction.

Since instructional activities do not always fit discretely within the modes of ICAP framework, the rubric was designed to allow for raters to use their best judgment in assigning ratings and doses, including hybrid rating of activities. The rubric enumerates guidelines, considerations, frequently asked questions, and examples for raters to use as reference.

Validity of the ICAP Instructional Rubric Scale.

Validity refers to the degree to which evidence supports that the interpretations of this instrument measure what is intended to be measured (Field, 2013). Since the ICAP framework has clear definitions that delimit the four constructs (i.e., interactive mode, constructive mode, active mode, and passive mode), in terms of construct validity, this instrument accurately represents the constructs measured. In terms of content validity, the ICAP instructional rubric rates the duration students were engaged in each of the four

ICAP modes. The instrument is representative of all aspects of the ICAP constructs since all aspects of instructional materials and course content can be rated by the ICAP framework. Finally, criterion validity for the ICAP framework has already been established based on results from studies (Chi et al., 2014; Chi et al., 2018; Wiggins et al., 2017) validating the ICAP's predictions for deeper student learning. This instrument also has evidence for criterion validity based on preliminary findings from a quasi-experimental study (Ha et al., 2022) that examined the relationship between the ICAP rated design of middle school mathematics instruction and student pre-post learning gains. Controlling for student pre-test scores, the study found that constructive dose had a relatively larger significant effect on student post-test scores than active dose.

Reliability of the ICAP Instructional Rubric Scale.

Within research, reliability refers to the repeatability or consistency (Trochim et al., 2016). In a prior study that utilized the ICAP instructional rubric instrument for rating the design of middle school mathematics classes, Ha et al. (2022) coded the data with an inter-rater reliability of Cohen's $\kappa = .877$.

Within this study, during the coding phase, the inter-rater and I selected approximately 50% of the data across the six iterations of the course to determine our inter-rater agreement. We initially coded approximately 5% of the data with an inter-rater agreement of about 70%. After reviewing the rubric's coding protocol and discussing the nuances of chemistry-based questions, we coded 50% of the data with an almost perfect inter-rater reliability agreement of Cohen's $\kappa = .991$. Our only discrepancy was in determining the proportion of questions or problems that belonged to the active or constructive category. We resolved any remaining disagreements via discussion.

Table 1
Examples of ICAP Rating

Assignment	Content	ICAP Rating
Homework	1) Which of the following is an element?	Active
	2) What is 875000 in scientific notation?	Active
	3) A gas-filled weather balloon with a volume of 65L is released at sea-level conditions of 745 torr and 25ºC. The balloon can expand to a maximum volume of 835L. When the balloon rises to an altitude at which the temperature is -5ºC and the pressure is 0.066atm, will it have expanded to its maximum volume?	Constructive
Recitation	1) Classify fluorine and sodium.	Active
	2) Rank the three types of bonds in order of strength.	Active
	3) Which of the samples release the most amount of heat upon cooling from 38ºC to 25ºC? Explain your answer.	Interactive
Post-lab	1) According to this lab, what form of electromagnetic radiation has energy slightly higher than that of visible light?	Active
	2) Which amount of salt caused the most ice to melt? Does it match the amount of salt predicted to melt the most ice? Explain.	Interactive
Investigation Lab Report	1) State learning goals for the investigation within the context of the purpose.	Active
	2) Write a one paragraph summary of the investigation; including the purpose of the investigation and the goals.	Interactive
Discussion	There are many chemistry informational websites. Identify at least one chemistry-related web resource that you found useful or at least entertaining. Explain what you like about your choice.	Constructive
Project	With your lab partners, present the findings of an investigation on a large poster board. The exposition need not be of a formal scientific type — feel free to pursue a more creative avenue for your presentation. Regardless of your presentation format, you must summarize your experiment and results in some way.	Interactive

Data Procedures

Since ICAP dose should be calculated based on the finest grain of data available, my first step was to identify what level of data was available consistently across courses. Instructor canvas course shells from previously completed courses included syllabi, course schedules, pre-recorded video lectures, homework, laboratory experiments, laboratory post-lab assignments, inquiry investigations, investigation lab reports, discussions, recitations, and a lab project. Since course materials were the finest grained data available, I based my ICAP dose ratings on the ratings of these individual materials. For reference, I included examples of ICAP ratings of items in instructional material in table 1.

The dose length of each material was determined in a few ways. Lectures, recitation, and laboratory experiments/inquiry investigations were determined by referring to the course catalog, syllabi, and schedules for the total time students spent on these activities during the course. The dose of pre-recorded lectures was determined by the video length. For online homework, the dose was calculated as the estimated time the online software suggested for completion. For homework from textbooks as well as post-lab assignments, investigation lab reports, the lab project, and discussion posts, I emailed instructors and asked them how long they expected for students to complete the assignments. Instructors generally reported a range of expected time for completion, and I determined the dose by averaging the minimum and maximum. For example, one instructor said that he expected homework to take about 1.5 to 2.5 hours depending on the student. Therefore, the average dose in this case was 2 hours per homework assignment.

CHAPTER 5

Results & Discussion of Findings

In this study, I analyze the data described in previous chapter using multiple regression. I model the relationship between student academic outcomes and course retention and course instruction as measured by ICAP rubric, controlling for student background characteristics and prior achievement. In this chapter, I present and discuss results of the estimation of the empirical models, each of which addresses one of the five research questions. Correspondingly, I estimate models with five dependent variables: participation level, final exam scores, course grades, course retention, and course attrition. I first present the description of participant demographics and baseline characteristics, and descriptive statistics for other variables in the models. I then provide the results of the regression analyses, and finally discuss and analyze the findings. All data analysis was conducted using IBM's SPSS Statistics software version 28.0.1.0 (142).

Descriptive Statistics

Participant Demographics and Characteristics

The participant data included the following demographic information: age, Pell grant status, whether they were a first-generation college student, sex, military service, part-time status, transfer status, and ethnicity. Student age ranged from 13 to 60 years of age, averaging 22.21 years, 63.4% of students identified their sex as female, 36.5% of students identified their sex as male, 3 students (0.1%) did not disclose their sex. 48.7% of students were Pell grant recipients, 37.4% of students were first generation college students, 3.0% of students identified as military service members, 28.8% of students were part time students, and 37.6% of students were transfer students. Race/ethnicity was categorized with the following breakdown: 2 or more ethnicities (5.9%), American Indian (1.8%), Asian American & Pacific Islander (AAPI: 9.4%), Black (5.3%), Hispanic (24.7%), White (49.2%), or no response (3.8%). In addition, the raw data for student race/ethnicity included two separate categories for Asian and Hawaiian/Pacific Islander. I chose to aggregate these two race/ethnicity categories for two reasons: a) much of the research literature around these two race/ethnicities use the singular demographic AAPI; and b) only a small sample 11 students reported their race/ethnicity as Hawaiian/Pacific Islander potentially skewing the demographic variable's relationship with the dependent variables in inferential analyses.

Student prior achievement data included incoming grade point average (GPA) and total scholastic aptitude test (SAT) scores. GPA and Total SAT provide measures of a student's baseline academic performance. Incoming GPA ranged from 1.04 to 4.79 points with a mean GPA of 3.22, including two subcategories for incoming high school GPA and incoming college transfer GPA. Incoming high school GPA averaged 3.44 and incoming transfer GPA averaged 2.93. Additionally, the raw data included both total SAT and American college testing (ACT) standardized test scores. I converted all ACT scores into equivalent SAT scores using the 2018 ACT/SAT concordance tables provided within the ACT website (American College Testing, 2018). SAT scores ranged from 790 to 1590 with a mean of 1162.1. The interquartile range for student SAT scores ranges from 1120 to 1360. In addition, the school reported that over two-thirds of incoming students have incoming GPAs of 3.42 or above with an average GPA of 3.50. The mean SAT score of students in this study falls within the lower end of the interquartile range and the mean incoming high school GPA was lower than the school's reported statistic, suggesting that participants in the study may be below average students in terms of measures of prior academic performance.

Table 2
Overall Descriptive Statistics for Participant Demographics

Nominal Variables	N	%	2019 (%)
Age			
Under 18	124	3.8	
18-24	2336	71.3	
25-29	440	13.6	
30-39	310	9.5	
40-49	55	1.7	
50+	12	0.3	
Pell Grant Recipient	1597	48.7	33
First Generation College Student	1225	37.4	30
Sex			
Male	1192	36.4	49.6
Female	2082	63.6	50.4
Not Disclosed	3	0.1	
Military Service	99	3.0	
Part-time Student	945	28.8	10.4
Transfer Student	1233	37.6	25.9
Race/Ethnicity			
2 or More	193	5.9	4.9
American Indian	58	1.8	1.5
Asian American & Pacific Islander	307	9.4	8.6
Black	175	5.3	4.7
Hispanic	808	24.7	26.5
White	1613	49.2	51.8
Not Disclosed	123	3.8	2.1

Interval Variables	N	Min	Max	Mean	Std. Dev.
Age	3277	13	60	22.2	5.9
Incoming GPA	2794	1.04	4.79	3.22	0.61
High school GPA	1580	1.04	4.79	3.44	0.55
Transfer GPA	1214	1.32	4.00	2.93	0.56
Total SAT Score	1561	710	1590	1162.1	166.1

Student race or ethnicity was for the most part tantamount to the average enrollee at the school, suggesting that student race or ethnicity did not affect their enrollment within this introductory chemistry course. However, there were several deviations of these participant demographics from the school's overall demographics. Students in the study were more likely to be Pell grant recipients, first generation college students, part-time students, and transfer students. A possible explanation for these demographic overrepresentations as well as lower prior academic performance, GPA and SAT, could be the introductory nature of this chemistry course. This introductory chemistry course was designed for students who did not take chemistry in high school. Additionally, students in the study were more likely to be female (63.6%). Possible explanations for why more females enrolled in this high school equivalent chemistry course could be attributed to two factors: a) gaps in STEM readiness for females (Card & Payne, 2021); and b) asynchronous online learners were more likely to be female (Wladis et al., 2015; Xu & Jaggars, 2011).

Dependent Variables

Student participation in the course, clicks per session, was measured by the average number of clicks students made in the LMS during a login session. The raw data from Canvas included total login sessions and total clicks for each student. I computed the clicks per session variable by dividing total clicks by total login sessions. My reasoning for using this variable in this manner was two-fold: a) total login sessions do not necessarily show the level of participation but rather just how many times students logged onto Canvas. There may be cases where students logged in many times but do not interact much with the LMS via clicking; and b) at the same time, the number of login sessions may indicate aspects of participation in terms of the frequency of accessing resources in the LMS. Clicks per session aggregates the total activity within LMS as a function of frequency of access and activity within each session, therefore acting as a variable that accounted for both aspects of LMS participation. On average, students clicked 5.66 times per login session with a standard deviation of 2.61 clicks per login session, see table 3.

Table 3
Overall Descriptives for Dependent Variables

Ratio Variables	**N**	**Min**	**Max**	**Mean**	**Std. Dev.**
Clicks Per Session (RQ 1)	2894	1.40	20.00	5.41	2.11
Final Exam Grade (RQ 2)	2740	1.00	112.00	69.05	19.79
# Weeks Participated (RQ 5)	235	1.00	15.00	6.20	3.11
Interval Variable	**N**	**Min**	**Max**	**Mean**	**Std. Dev.**
Course Grade (RQ 3)	2888	0	4.33	3.17	1.17
Nominal Variable				**N**	**%**
Course Withdrawal (RQ 4)				235	7.2

Final exam grade was the indicator of student testing performance. Course exam progression was designed in one of two ways: a) several midterm exams with a culminating final exam; or b) several midterm exams only. For courses that opted for the latter option, the midterm exams were more substantive since there was no final exam. Since courses utilizing the latter option do not have a true final exam, I average student scores for all midterm exams as a proxy for student final exam grades. The mean final exam grade for the course was 69.05 out of a 100 with a standard deviation of 19.79.

Course grade was the indicator for overall course performance. The raw data included course letter grades from F to A+. Grade point conversions were calculated as follows: F = 0.00; D = 1.00; C- = 1.67; C = 2.00; C+ = 2.33; B- = 2.67; B = 3.00; B+ = 3.33; A- = 3.67; A = 4.00; A+ = 4.33. Course grade was an interval variable, since a GPA of 0.00 is not a true zero, representing any grade under 65%. Therefore, the grades were not true ratios of grade value percentages in the course. The mean course grade was 3.17, between a letter grade of B and B+, with a standard deviation of 1.17 grade points.

The course retention variable indicated whether a student withdrew from the course. The information on course withdrawal for all students was included in the raw data. Since the initial enrollment period allows for students to drop courses without penalty for two weeks, I removed students who withdrew within the first two weeks of the course from the variable as these students most likely withdrew for personal or logistical reasons and that decision was not associated with the course design or delivery. Excluding students who dropped from the course, 235 (7.2%) students withdrew from the course.

Course attrition was measured by the number of weeks participated for students who withdrew from the course, indicating how long students participated in the course before withdrawing. Depending on the semester when the course was offered, courses had a maximum duration of either 14 or 15 weeks. This variable was computed based on the total number of weeks students participated via login sessions and clicks. On average, students who withdrew stayed in the course for an average of 6.2 weeks with a standard deviation of 3.11 weeks. For clarification on the frequency, students who only participated for only 1 or 2 weeks were also included if they withdrew from the course after the second week, indicating that some students did not log into canvas until at least a week into the course.

Independent Variables

The primary independent variables for this study were the ICAP Instructional Rubric doses that enumerates the total number of minutes students were engaged in each mode: passive dose, active dose, constructive dose, and interactive dose. Values for these variables were computed by the researcher and inter-rater based on the instructions of the ICAP instructional rubric. Essentially, each course activity was assigned an ICAP dose value based on the percentage of ICAP modes of engagement of instructional materials or activities multiplied by the number of minutes students engaged in each activity. The assigned ICAP dose values were the sum of ICAP doses for all course activities within each course.

Table 4
Overall Descriptives for Independent Variables

Ratio Variables	**N**	**Min**	**Max**	**Mean**	**Std. Dev.**
Passive Dose	3277	1219	1950	1630.2	334.5
Active Dose	3277	1912	3027	2490.5	501.1
excluding labs	3277	792	1884	1299.5	432.3
excluding labs/recitations	3277	792	1397	947.3	226.9
Constructive Dose	3277	0	470	228.1	187.8
excluding labs	3277	0	342	173.3	149.0
excluding labs/discussions	3277	0	342	75.4	119.3
homework only	3277	0	222	60.7	89.2
discussions only	3277	0	330	97.9	150.8
Interactive Dose	3277	0	1527	344.7	483.8
excluding labs	3277	0	215	28.9	51.9
Nominal Variables				**N**	**%**
Mode of instruction					
Face to Face (F2F)				788	24.0
Asynchronous (AOI)				1282	39.1
Blended Synchronous (BSL)				1207	36.8
AOI or BSL				2489	76.0

ICAP dose also included measured variations that excluded different instructional activities. These variations were computed since the dependent variables vary in terms of relevance to instructional materials or activities. For example, final exam questions did not effectively measure the kinds of learning that occurred during laboratory experiments, lab reports, lab research poster sessions, etc. In the following section on inferential statistics, I give reasoning for utilizing different variations of ICAP doses within each model.

In terms of ICAP ratings for different course activities, lectures were rated as engaging students in the passive mode since students were expected to pay attention and not necessarily expected to take notes. There were three modes of lectures in the data: pre-recorded videos for AOI courses, Zoom lectures for BSL courses, and classroom lectures for F2F courses.

Homework was rated as engaging students in either the active or constructive modes; basic or easier homework problems such as concept recall and skill practice were rated as active-level and more difficult problems in context that required students to model and generate ideas were rated as constructive-level. There were two methods of delivery for homework: proprietary online software or selected textbook questions. Online software-based homework tended to ask fewer constructive questions.

Recitation was rated in a similar fashion as homework. The only major difference was that items that would have been rated as constructive for homework were rated as interactive since students collaboratively solved recitation problems.

Laboratory had five different components: experiments, post-lab questions, investigations, investigation lab reports, a lab research poster project. Procedural labs were considered experiments, paired with post-lab questions whereas inquiry-based labs were considered investigations, paired with investigation lab reports. Experiments were rated as active since students conducted labs by following lists of instructions and procedures whereas investigations were classified as interactive since students collaboratively solved an inquiry lab problem using a problem-based learning approach. Investigations posed a related chemistry question without explicit instructions for how to design and carry out an experiment to answer the question. This necessitated the collaborative generation of an experimental design and procedure, which engaged students in the interactive mode. Post-lab assignments were rated as either active or interactive depending on whether each question required students to collaboratively generate ideas. Generative questions were classified as interactive since students conducted experiments and post-labs collaboratively. Investigation lab reports were rated as either active or interactive based on the guidelines for investigation reports. Portions of the lab write-up were generative such as writing discussions and conclusions whereas some portions were not generative such as reporting data that was collected during the investigation. Lab research poster projects were rated as interactive since students were either presenting their poster or formulating and asking questions regarding other students' posters.

Discussion posts were rated as constructive since the posts required students to summarize and elaborate on discussion questions individually.

Other independent variables included in the models were the indicators of the mode of instruction: F2F, AOI, and BSL. At the university, courses that were normally F2F were converted into BSL courses from the spring of 2020 to spring of 2021 due to the COVID-19 pandemic. In addition, I computed a new variable that combined AOI and BSL courses since these courses were very similar in how they were delivered. Within this study, the only major difference between BSL and AOI was whether lectures were delivered via pre-recorded videos or delivered synchronously via Zoom conferencing. All other aspects of BSL courses were conducted online in the LMS. While there were no recordings of synchronous lectures within the LMS, this study assumed that the adapted synchronous chemistry lectures did not differ much from their F2F counterparts; the courses most likely did not include much interaction between the lecturer and students since these introductory courses were very large with an average class section size of 252 students.

While synchronous and asynchronous lectures may differ in terms of student motivation and self-regulation, these instructional activities were both classified by the ICAP framework as engaging students in the passive mode. From the standpoint of cognitive engagement, pre-recorded video and Zoom lectures were essentially the same. Although students were able to pause, rewind, and revisit pre-recorded videos, repeated viewing of pre-recorded lectures potentially only increases passive dose. Since the delivery of AOI and BSL were essentially the same in terms of the ICAP framework, keeping these discrete categories introduced noise in the data that could lead to high levels of variance inflation due to multicollinearity. For these reasons, I decided to combine AOI and BSL variables into the category: AOI or BSL.

Results

Multiple regression analysis was used to determine the relationships between course ICAP dose ratings and dependent measures, including student participation, final exam grades, course grades, course retention, and course attrition. For each of these five analyses, I ran multiple models to ascertain the final model. In the first model, I included the primary predictors of this study, ICAP dose. The second model included the mode of instruction, and the third model included indicators of prior academic success, incoming GPA and total SAT scores. For the fourth model, I also included student demographic variables. Some analyses included a fifth model that reduced the number of independent variables with explanations for the reduced model.

I present the results below, organized by the dependent variable, an explanation of student sample and inclusion/exclusion criteria for independent variables, and model progression. All subsequent linear tests supported assumptions of linearity, normality, and homoscedasticity. In addition, variance inflation factor (VIF) values for the final model were below 10, indicating no symptoms of multicollinearity. Furthermore, I adopted a conservative approach on variance inflation; independent variables with VIF values above 5 were removed from subsequent models. Tests were conducted at significance level of $\alpha = 0.05$.

For clarity of reporting relationships, I discuss the findings for each dependent variable using standardized beta coefficients for linear regressions and odds ratios for logistic regressions. For linear regression, the standardized beta coefficients represent the relationship between each independent and dependent variables in terms of standard deviations. In the discussions for each dependent variable, changes in the dependent variables associated with the changes in independent variables are calculated by multiplying the standardized coefficient β of the independent variable with the standard deviation (SD) of the DV, $\beta_{IV} * SD_{DV}$.

Table 5
Course Participation

	Model 1 - β	Model 2 - β	Model 3 - β	Model 4 - β	Model 5 - β
R / Adjusted R^2 / ANOVA Significance	.714/.509***	.724/.523***	.712/.504***	.723/.516***	.731/.531***
Passive Dose	-.697***	-.615***	-.609***	-.609***	-.651***
Active Dose - excluding labs/recitations	.069***	.020	.031	.033	.000
Constructive Dose - excluding labs	.140***	.131***	.128***	.142***	
Constructive Dose - homework only					.139***
Constructive Dose - discussions only					.093***
AOI or BSL		.137***	.170***	.172***	.141***
Total SAT			-.095***	-.083***	
Incoming GPA			.038		
Age				-.045	-.061***
Pell Grant				-.003	-.004
First Generation				.006	.020
Female				.068***	.066***
Military Service				.007	.024
Part-time				.026	.030
Transfer				.014	.040*
2 or more Ethnicities				.001	.003
American Indian				.002	.013
Asian American & Pacific Islander				.056**	.045***
Black				-.001	-.013
Hispanic				.030	.024

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$;

These values are the associated increases and decreases in the DV for relative unit changes in SD_{IV}. For logistic regression, the odds ratios represent the relationship between each IV and DV based on likelihood ratios. In the discussion for odds ratios, percent likelihood was calculated as $e^{\beta} * 100\%$. Increases or decreases of % likelihood was calculated as $(1 - e^{\beta}) * 100\%$.

Course Participation

Course participation was analyzed as a function of the number of student clicks within the LMS per LMS login session. Since students who dropped or withdrew probably have different participation patterns, I excluded those students from this analysis, including only students who completed the full duration of the courses. Data examining course participation are presented in table 5.

For this dependent variable, I excluded recitation and laboratory for two reasons: a) students in F2F versions of the course completed these course components in person rather than in the LMS; and b) laboratory and recitation do not directly relate to course LMS usage. I included measures of remaining course components including lectures, homework, and discussion posts. Although lectures for students in F2F course sections were completed in person, lectures relate to Canvas participation since lecture resources were housed within the LMS as well as supplemental resources that provide further enrichment based on lectures. Since none of these remaining activities included interactive ratings, interactive dose was excluded from this analysis.

Model 1. The first model's results indicated that student clicks per session were significantly predicted by an overall model including ICAP dose [$R = .714$, $R^2_{adj} = .509$, $p < .001$]. This model accounted for 50.9% of the variance in the student participation.

The regression coefficients, β, indicated that passive dose had a significant negative relationship with student participation (β = -.697, p < .001), active dose had a significant positive relationship with student participation (β = .069, p < .001), and constructive dose had a significant positive relationship with student participation (β = .140, p < .001).

Model 2. The second model included mode of instruction, AOI or BSL. Results indicated that clicks per session were significantly predicted by the updated model that included mode of instruction [R = .724, R^2_{adj} = .523, p < .001]. This model accounted for 52.3% of the variance in student participation. The regression coefficient for AOI or BSL indicated that it had a significant positive relationship with student participation (β = .137, p < .001). In addition, controlling for the mode of instruction, active dose did not have a significant relationship with course participation.

Model 3. The third model included student total SAT scores and incoming GPA. Results indicated that clicks per session were significantly predicted by the updated model that included incoming GPA [R = .712, R^2_{adj} = .504, p < .001]. This model accounted for 50.4% of the variance in student participation. The regression coefficient for total SAT score indicated that it had a significant negative relationship with student participation (β = -.095, p < .001). Incoming GPA did not significantly predict student participation (β = .038, p > .05). Therefore, I removed the student incoming GPA variable from subsequent models.

Model 4. The fourth model included student demographics in addition to all other independent variables. Results indicated that clicks per session were significantly predicted by the updated model that included student demographics [R = .723, R^2_{adj} = .516, p < .001]. This model accounted for 51.6% of the variance in student participation. The regression coefficients indicated that student sex - female (β = .068, p < .001) and race/ethnicity - AAPI (β = .056, p < .01) had significant positive relationships with student participation as compared to students who identified as white.

Model 5. Considering discussions posts were expected to relate with course participation, I ran a fifth model that specified the constructive dose for homework or discussions to elucidate the relative effect of constructive - homework dose on course participation. In addition, since the data only included total SAT scores for approximately half of the students, see table 6, I chose to remove total SAT scores from the final model. Results indicated that clicks per session were significantly predicted by the updated model that included student demographics [$R = .731$, $R^2_{adj} = .531$, $p < .001$]. This model accounted for 53.1% of the variance in student participation. The regression coefficients for this model indicated that constructive dose - homework only (β = .139, p < .001), constructive dose - discussion only (β = .093, p < .001), mode of instruction - AOI or BSL (β = .141, p < .001), student sex - female (β = .066, p < .001), transfer status (β = .040, p < .05), and race/ethnicity - AAPI (β = .045, p < .001) had significant positive relationships with student participation. Passive dose (β = -.651, p < .001) and age (β = -.061, p < .001) had significant negative relationships with student participation. Active dose, Pell grant recipient, first generation status, part-time status, transfer status, military service, race/ethnicity - 2 or more, American Indian, Black, and Hispanic did not have significant relationships with course participation.

Discussion of Findings for Course Participation. For every 330.8 minutes, or roughly 5.5 hours, of passive dose for homework, students likely performed 1.37 fewer

Table 6

Course Participation - Descriptive Statistics - Significantly Related Variables

Model 5	N*	Mean	SD	β	β * SD_{DV}
Click per Session	2894	5.41	2.11		
Passive Dose	2894	1648.4	330.8	-.651	-1.37
Constructive Dose - HW	2894	64.9	91.0	.139	.29
Constructive Dose - Discussion	2894	93.6	148.8	.093	.20
AOI or BSL	2188			.141	.30
Total SAT	1424	1167.1	166.5	-.083	-.18
Age	2894	22.0	5.8	-.061	-.13
Female	1805			.066	.14
Transfer	1010			.040	.08
AAPI	274			.045	.09

*excluding students who dropped or withdrew and including only students with clicks per session

clicks per login session, see table 6. Although passive dose was found to have a significant negative effect on course participation in the course LMS, the variance within the passive dose in terms of mode of instruction explains this result. AOI courses utilized pre-recorded video lectures whereas F2F and BSL courses had lectures that spanned a set amount of time each semester. There was not much variation for lecture seat time with F2F and BSL whereas AOI pre-recorded lectures tended to be shorter than their live lecture counterparts. Further analyses of descriptive data revealed that the average total lecture length for F2F and BSL courses was 1894.4 minutes whereas the average total length of pre-recorded videos for AOI was 1219.0 minutes. Therefore, the negative passive dose relationship with course participation was expected since students in AOI courses were more likely to be active in Canvas to click on resources and pre-recorded lectures.

Constructive dose was found to have a significant positive effect on course participation in the course LMS. Within this model, constructive dose included homework assignment questions that were rated as constructive, as in problems that required students to use generative learning strategies, and discussion posts. These results imply that students who take courses with an increased portion of assignments requiring more generative learning strategies may participate more in the course LMS. For every 91 minutes, or roughly 1.5 hours, of constructive dose for homework, students likely performed 0.29 more clicks per login session. For every 148.8 minutes, or roughly 2.5 hours, students spent on discussion posts, students likely performed 0.20 more clicks per login session. In terms of relative effect of time, constructive level homework problems had a 137% greater effect or more than twice the efficacy on student participation in the LMS than discussions. This value was calculated as the ratio of homework and discussion in terms of clicks per session per dosage [(0.29 clicks per session / 91.0 homework dose) / (0.20 clicks per session / 148.8 discussion dose) = 2.37].

A possible explanation for changes in participation associated with ICAP doses of homework could be that a student solving more difficult constructive level problem-solving questions may be more likely to utilize resources available in Canvas whereas a student solving easier active level homework questions may be less likely to utilize LMS resources. Increased participation in the LMS was expected with the inclusion of discussion posts since students primarily read and construct responses within Canvas.

While the effect on the number of clicks students perform per login session may be modest at best, these findings could still be very useful for looking at larger trends within large data sets. In addition, these modest changes may have also had an effect on student course performance. I performed a linear regression between course participation and course grades and found a significant positive relationship between the variables ($R = .264$, $R^2 = .07$), which explained approximately 7% of the variance in course grades.

Courses that were AOI or BSL were associated with greater participation in the LMS with 0.30 more student clicks per session than F2F courses. This was an expected result since students rely on the LMS more for online courses than for F2F courses.

In model 4, total SAT score had a negative significant relationship with course participation. For every increase of 166.5 total SAT score, students were less likely to participate in the LMS with 0.18 fewer clicks per login session. A possible explanation for these results is that students who enter the course with greater SAT scores may be higher performing students who rely less on resources in the LMS for completing homework.

Students who identified as female were more likely to participate in the LMS with 0.14 more clicks per login session than those who identified as male. A prior case study showed women were more likely to participate in the LMS per login session than males (Morante et al., 2017). A possible explanation for these results is that females may be more inclined to participate more in LMS due to documented higher levels of conscientiousness (Feingold, 1994; Costa et al., 2001), which are associated with an increase in student performance (Richardson & Abraham, 2009). Conscientiousness is a personality trait associated with responsibility, persistence, trustworthiness, and dependability (Conrad & Patry, 2012). These results provide further evidence that female students may be more likely to participate in LMSs than male students.

Table 7
Final Exam Grade

	Model 1 - β	Model 2 - β	Model 3 - β	Model 4 - β	Model 5 - β
R / Adjusted R^2 / ANOVA Significance	.460/.210***	.508/.258***	.628/.391***	.637/.396***	.578/.329***
Passive Dose	.827^{***+}	.510***	.338***	.372***	.361***
Active Dose - excluding labs	-.599^{***+}				
Constructive Dose - excluding labs/disc	.170***	.030	.124***	.130***	.084***
Interactive Dose - excluding labs	.156***	.022	.039	.024**	.035***
AOI or BSL		.273***	.348***	.340***	.284***
SAT Total			.290***	.284***	
Incoming GPA			.163***	.189***	.227***
Age				.021	.010
Pell Grant				-.008	-.049*
First Generation				.024	-.035
Female				-.070**	-.088***
Military Service				.018	.012
Part-time				-.018	-.055**
Transfer				.062	-.019
2 or more Ethnicities				-.008	.001
American Indian				-.035	-.040*
Asian American & Pacific Islander				.011	-.005
Black				-.021	-.016
Hispanic				-.036	-.069***

$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$;

Transfer students were also more likely to participate in the LMS with 0.08 more clicks per login session than those who were not transfer students. These results imply that there may be something categorically different about transfer students in terms of their participation patterns in LMS courses. These results can be potentially explained by further analyses of descriptives which showed that of 1010 transfer students, 663 (65.6%) identified as female, who were more likely to participate in LMSs.

Students who identified as AAPI were more likely to participate in the LMS with an increase of 0.09 clicks per session as compared to students who identified as white. Prior research has established that, in general, the majority of AAPI perform better than white students in college (Fischer, 2007). In addition, prior research has shown that AAPI students were more extrinsically motivated than white students in terms of fear of failure (D'Lima et al., 2014). These factors for AAPI students may help explain why AAPI students participated more in LMS in the study.

Final Exam Grade

Analysis of models for final exam grade included the full sample. At the same time, since only students who completed the full duration of the courses had final exam grade scores, the model effectively excluded students who dropped or withdrew from the course. Data examining final exam grades are presented in table 7.

For this dependent variable, I excluded laboratory since all components of laboratory were not directly assessed by final examinations. Rather, the laboratory course components assess student skills such as ability to conduct experiments, writing lab reports, and answering lab-specific questions. Moreover, some courses included a separate final examination for laboratory. I also excluded discussions since discussion

Table 8
Correlation Matrix for ICAP Dose – Final Exam Grade

		Passive	Active	Constructive	Interactive
Passive	R	1	.957**	.598**	.510**
	Sig.		.000	<.001	<.001
Active	R	.957**	1	.751**	.607**
	Sig.	.000		.000	<.001
Constructive	R	.598**	.751**	1	.573**
	Sig.	<.001	.000		<.001
Interactive	R	.510**	.607**	.573**	1
	Sig.	<.001	<.001	<.001	
	N	2740	2740	2740	2740

**. Correlation is significant at the 0.01 level (2-tailed).

posts did not bear relevance to the questions in the final exam. I included measures for the other course components that were relevant for performance on final exams, including lectures, homework, and recitation.

Model 1. The first model's results indicated that final exam grades were significantly predicted by an overall model including ICAP dose [$R = .460$, $R^2_{adj} = .210$, $p < .001$]. This model accounted for 21% of the variance in final exam grades. The regression coefficients, β, indicated that passive ($\beta = .817$, $p < .001$), constructive ($\beta = .170$, $p < .001$), and interactive ($\beta = .156$, $p < .001$) doses had significant positive relationships with final exam grades and active dose had a significant negative relationship with final exam grades ($\beta = -.599$, $p < .001$).

At the same time, the VIF for passive and active dose were both greater than a value of 5. To elucidate the problem of variance inflation, I analyzed the correlation matrix among the ICAP doses that included all course components, see table 8. Results show that passive dose and active dose were highly correlated, $R = .957$. In addition, passive dose was moderately correlated with both constructive, R = .598, and interactive doses, R = .510. Active dose was highly correlated with constructive dose, R = .751, and moderately correlated with interactive dose, R = .573. I chose to remove active dose from this model since it had the highest correlation among the ICAP doses. Removal of active dose resolved the issue of VIF, and subsequent models no longer had any VIF values above 5.

Model 2. The second model included mode of instruction, AOI or BSL, into the model. Results indicated that final exam grades were significantly predicted by the updated model that included mode of instruction [R = .508, R^2_{adj} = .258, p < .001]. This model accounted for 25.8% of the variance in final exam grades. The regression coefficient for AOI or BSL indicated that it had a significant positive relationship with final exam grades (β = .273, p < .001). Removing active dose and controlling for mode of instruction, both constructive and interactive doses did not have a significant relationship with final exam grades within this model.

Model 3. The third model included student total SAT scores and incoming GPA. Results indicated that final exam grades were significantly predicted by the updated model that included total SAT scores [R = .628, R^2_{adj} = .391, p < .001]. This model accounted for 39.1% of the variation for final exam grades. The regression coefficients for SAT total (β = .290, p < .001) and incoming GPA (β = .163, p < .001) indicated that they had significant positive relationships with final exam grades. Controlling for markers of prior academic performance, constructive dose (β = .124, p < .001) had a significant positive relationship with final exam grades within this model.

Model 4. In addition to all previous variables, the fourth model also included student demographics. Results indicated that final exam grades were significantly predicted by the updated model that included student demographics [R = .637, R^2_{adj} = .396, p < .001]. This model accounted for 39.6% of the variance in final exam grades. The regression coefficients for this model indicated that student sex - female (β = -.070, p < .01) had a significant negative relationship with final exam grades. This result did not indicate that sex had a relationship with exam grades but rather indicated that on average, students who identified as female within this course scored fewer points on the final exam than those who identified as male. Controlling for student demographics, interactive dose had a significant positive relationship with final grades within this model (β = .024, p < .01).

Model 5. I removed the total SAT score from the fifth model since the data only included total SAT scores for approximately half of the students. Results indicated that final exam grades were significantly predicted by the updated model that included student demographics [R = .578, R^2_{adj} = .329, p < .001]. This model accounted for 32.9% of the variance in final exam grades. The regression coefficients for this model indicated that passive dose (β = .361, p < .001), constructive dose (β = .084, p < .001), interactive dose (β = .035, p < .01), AOI or BSL (β = .284, p < .001), and incoming GPA (β = .227, p < .001) had significant positive relationships with final exam grades. In addition, pell grant recipients (β = -.049, p < .05), student sex - female (β = -.088, p < .001), part-time status (β = -.055, p < .01), race/ethnicity - American Indian (β = -.040, p < .05), and Hispanic (β = -.069, p < .001) had significant negative relationships with final exam grades. Age, first-generation status, military service, transfer status, and race/ethnicity -

Table 9

Final Exam Grades - Descriptive Statistics - Significantly Related Variables

Model 5	N*	Mean	SD	β	β * SD_{DV}
Final Exam Grade	2740	69.05	19.79		
Passive Dose	2740	1659.8	327.4	.361	7.14
Constructive Dose	2740	83.4	123.8	.084	1.66
Interactive Dose	2740	31.4	53.3	.035	.69
AOI or BSL	2056			.284	5.62
SAT Total (Model 4)	1362	1169.25	166.16	.284	5.62
Incoming GPA	2300	3.28	.60	.227	4.49
Pell Grant	1296			-.049	-.97
Female	1705			-.088	-1.74
Part-time	566			-.055	-1.09
American Indian	45			-.040	-.79
Hispanic	680			-.069	-1.37

*including only students who received a final exam grade

2 or more, AAPI and Black as compared to that of students who identified as white did not have significant relationships with final exam grades

Discussion of Findings for Final Exam Grades. Passive dose was found to have a significant positive effect on final exam grades. For every 327.4 minutes, or roughly 5.5 hours, of passive dose, students likely scored 7.14 points higher on the final exam, see table 9. As discussed in the previous section, passive dose varied as a function of whether the mode of instruction of courses were AOI since BSL and F2F courses had the same passive dose due to set lecture times. A reasonable explanation for these results is that students who took the AOI version of the course were different from the population of students who take BSL and F2F. Further analysis of descriptives for incoming GPA of AOI students revealed that students in AOI sections of the course had a mean incoming GPA of 2.92, SD = 0.6, whereas students in either F2F or BSL sections of the course had a mean incoming GPA of 3.44, SD = 0.5. Therefore, the variance in final exam grades may be better explained by sample differences rather than variance in passive dose. Another explanation for this effect is that some of the final exam questions likely contained a large portion of recall problems that aligned with student learning gains from passive and active dose, which were the modes that students were largely engaged in.

Constructive dose was found to have a significant positive effect on final exam grades. For every 123.8 minutes, or roughly 2 hours, students spent on constructive questions in homework, students likely scored 1.66 more points on the final exam. Within this model, constructive dose included homework assignment questions that were rated as constructive. These results imply that students who take courses with an increased portion of homework requiring more generative learning strategies may perform better on final exams. A likely explanation for this potent effect on final exam grades is that solving more difficult problem-solving questions in homework prepares students for solving equivalent problems on final exams.

Interactive dose was found to have a significant positive effect on final exam grades. For every 53.3 minutes, or roughly 1 hour, spent on interactive recitation problems, students likely scored 0.69 more points on the final exam. These results provide further evidence for generative problem-solving questions improving student final exam grades. Accounting for variance in dose length ratio, the relative effect of interactive dose on final exam grades was approximately equivalent to that of constructive dose with only a 0.1% difference in favor of the interactive mode. This suggests that interactive recitation problems may be just as effective for improving student final exam grades as constructive homework problems. At the same time, the effect of interactive dose was likely underestimated. The data available in this study did not ascertain if students were equitably co-constructing knowledge while solving generative problems in recitation. If students in groups unevenly distributed workload or simply divided responsibility for the recitation, the effective average generative dose for each student would be less than the reported interactive dose. These results provide further evidence suggesting that the ICAP framework's predictions for student learning are valid. Considering this study did not ascertain the quality of collaboration, the interactive dose was at a minimum likely tantamount to constructive dose in effect and possible more effective if students did not collaborate equitably and co-generatively.

Courses that were AOI or BSL were associated with greater final exam grades with students likely scoring 5.62 more points on the final exam. These results were expected and explained by prior research that shows that students perform better on online courses due to more accessible cheating than in F2F classes with proctored examinations (Bilen & Matros, 2021; Watson & Sottile, 2010). Another explanation for improved test scores is that students experienced less text anxiety for online tests and therefore performed better (Stowell & Bennett, 2010).

Total SAT score within model 4 was associated with higher scores on final exam grades with students likely scoring 5.62 more points on the final exam for every 166.16 points increase in total SAT score. In addition, for every 0.6 increase in incoming GPA, students were likely to score 4.49 more points on the final exam. These results were expected since students who scored higher on total SAT scores and have higher incoming GPA were likely to be higher performing students in college and more likely to be comfortable with test-taking as well as familiar with effective test-taking strategies.

Table 10
Course Grade

	Model 1 - β	Model 2 - β	Model 3 - β	Model 4 - β	Model 5 - β
R / Adjusted R^2 / ANOVA Sig.	$.241/.057^{***}$	$.270/.072^{***}$	$.489/.235^{***}$	$.522/.261^{***}$	$.459/.205^{***}$
Passive Dose	$-.461^{***+}$				
Active Dose	$.575^{***+}$	$.281^{***}$	$.233^{***}$	$.267^{***}$	$.212^{***}$
Constructive Dose	$.121^{***}$	$.121^{***}$	$.219^{***}$	$.215^{***}$	$.169^{***}$
Interactive Dose	$.302^{***}$	$.116^{***}$	$.134^{***}$	$.155^{***}$	$.108^{***}$
AOI or BSL		$.169^{***}$	$.215^{***}$	$.198^{***}$	$.187^{***}$
Total SAT			$.218^{***}$	$.200^{***}$	
Incoming GPA			$.272^{***}$	$.300^{***}$	$.314^{***}$
Age				$.098^{**}$	$.116^{***}$
Pell Grant				-.030	$-.046^{*}$
First Generation				-.008	$-.050^{*}$
Female				-.031	-.034
Military Service				$.093^{***}$	$.086^{***}$
Part-time				$-.131^{***}$	$-.117^{***}$
Transfer				-.057	-.036
2 or more Ethnicities				-.047	-.013
American Indian				-.047	$-.040^{*}$
AAPI				-.001	-.011
Black				-.020	$-.050^{**}$
Hispanic				-.047	$-.073^{***}$

+VIF > 5; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$;

Pell grant recipients scored on average 0.97 fewer points on the final exam. Considering receiving a Pell grant is a proxy for lower socioeconomic status, these results echo prior research that has shown that lower SES students typically attain lower GPA in college (An, 2013).

Students who identified as female likely scored 1.74 fewer points on the final exam than that of their male counterparts. These results were in line with prior research that has shown that females were less likely to be STEM ready than males (Card & Payne, 2021). In addition, studies have shown that females perform more poorly on examinations in STEM courses (Ballen et al., 2017; Eddy et al., 2014; Matz et al., 2017).

Part-time students likely scored 1.09 fewer points on the final exam than full-time students. Given part-time students often have other responsibilities such as parenthood and careers (Cohen & Greenberg, 2011), a plausible reason could be that part-time students might have other priorities in life that may make performing well on final exams more difficult.

As compared to students who identified as white, American Indian students were likely to score 0.79 fewer points and Hispanic students were likely to score 1.37 fewer points on the final exam. These findings are in line with prior research which has established that students who identify as American Indian or Hispanic typically earn lower grades in college (Perie & Moran, 2005).

Course Grade

Analysis of models for course grades included the full sample. At the same time, since only students who completed the full duration of the courses received a course grade, the model effectively excluded students who dropped or withdrew from the course.

Data examining course grades are presented in table 10. For this DV, I included all instructional activities for ICAP dose since course grades relate to all aspects of the course. I included measures of course components including lectures, homework, recitation, laboratory, and discussion posts.

Model 1. The first model's results indicated that student course grades were significantly predicted by an overall model including ICAP dose [$R = .241$, $R^2_{adj} = .057$, $p < .001$]. This model accounted for 5.7% of the variance in course grades. The regression coefficients, β, indicated that active dose ($\beta = .575$, $p < .001$), constructive dose ($\beta = .121$, $p < .001$), and interactive dose ($\beta = .302$, $p > .001$) had positive significant relationships with course grades had a significant positive relationship with course grades. In addition, passive dose ($\beta = -.461$, $p > .001$) had a significant negative relationship with course grades.

Table 11
Correlation Matrix for ICAP Dose – Course Grade

		Passive	Active	Constructive	Interactive
Passive	R	1	.879**	-.611**	.648**
	Sig.		.000	<.001	.000
Active	R	.879**	1	-.808**	.273**
	Sig.	.000		.000	<.001
Constructive	R	-.611**	-.808**	1	.065**
	Sig.	.000	.000		<.001
Interactive	R	.648**	.273**	.065**	1
	Sig.	.000	<.001	<.001	
	N	2888	2888	2888	2888

**. Correlation is significant at the 0.01 level (2-tailed).

However, within this model, the VIF was greater than a value of 5 for passive and active doses. To elucidate the problem of variance inflation, I analyzed the correlation matrix among the ICAP doses that included all course components, see table 11. Results show that passive dose and active dose were highly correlated, R = .879. In addition, passive dose was moderately correlated with both constructive, R = -.611, and interactive doses, R = .648. Active dose was highly correlated with constructive dose, R = -.808, and had low correlation with interactive dose, R = .273. I chose to remove passive dose from this model since passive dose had the greater variance inflation than active dose among all ICAP doses. Removal of passive dose resolved the issue of VIF, and subsequent models no longer had any VIF values above 5.

Model 2. The second model included mode of instruction, AOI or BSL, into the model. Results indicated that course grades were significantly predicted by the updated model that included mode of instruction [$R = .270$, $R^2_{adj} = .072$, $p < .001$]. This model accounted for 7.2% of the variance in course grades. The regression coefficient for AOI or BSL indicated that it had a significant positive relationship with course grades ($\beta = .169$, $p < .001$).

Model 3. The third model included total SAT score and incoming GPA. Results indicated that course grades were significantly predicted by the updated model that included total SAT and incoming GPA [$R = .489$, $R^2_{adj} = .235$, $p < .001$]. This model accounted for 23.5% of the variance in course grades. The regression coefficients, β, indicated that both total SAT ($\beta = .218$, $p < .001$) and incoming GPA ($\beta = .272$, $p < .001$) had significant positive relationships with course grades.

Model 4. The fourth model included student demographics. Results indicated that course grades were significantly predicted by the updated model that included student demographics [$R = .522$, $R^2_{adj} = .261$, $p < .001$]. This model accounted for 26.1% of the variance in course grades. The regression coefficients for the model indicated that age ($\beta = .098$, $p < .01$) and military service ($\beta = .093$, $p < .001$) had significant positive relationships with course grades. Part-time status ($\beta = -.131$, $p < .001$) had a significant negative relationship with course grades.

Model 5. The fifth model removed the total SAT score since the data only included total SAT scores for approximately half of the students. Results indicated that course grades were significantly predicted by the updated model that removed total SAT scores [$R = .459$, $R^2_{adj} = .205$, $p < .001$]. This model accounted for 20.5% of the variance in course grades. The regression coefficients for the final model indicated that active dose ($\beta = .212$, $p < .001$), constructive dose ($\beta = .169$, $p < .001$), interactive dose ($\beta = .108$, $p < .001$), mode of instruction - AOI or BSL ($\beta = .187$, $p < .001$), incoming GPA ($\beta = .314$, $p < .001$), age ($\beta = .116$, $p < .001$), and military status ($\beta = .086$, $p < .001$) had significant positive relationships with course grades. Pell grant recipients ($\beta = -.046$, $p < .05$), first generation college students ($\beta = -.050$, $p < .05$), and part-time status ($\beta = -.117$, $p < .001$) as well as race/ethnicity - American Indian ($\beta = -.040$, $p < .05$), Black ($\beta = -.050$, $p < .01$), and Hispanic ($\beta = -.073$, $p < .001$) had significant negative relationships with course grades. Sex - female, transfer status, race/ethnicity - 2 or more, and AAPI did not have a significant relationship with course grades.

Discussion of Findings for Course Grades. Active dose was found to have a significant positive effect on course grades. For every 495.7 minutes, or roughly 8 hours, of active dose for, students on average attained 0.25 grade points for the course grade, see table 12. Within this model, active dose included homework, recitation, and laboratory. Constructive dose was found to have a significant positive effect on course grades as

Table 12

Course Grade - Descriptive Statistics - Significantly Related Variables

Model 5	N*	Mean	SD	β	β * SD_{DV}
Course Grade	2888	3.17	1.17		
Active Dose	2888	2512.2	495.7	.212	.25
Constructive Dose	2888	225.3	186.7	.169	.20
Interactive Dose	2888	368.8	495.1	.108	.13
AOI or BSL	2184			.187	.22
Total SAT (Model 4)	1424	1167.1	166.48	.200	.23
Incoming GPA	2432	3.26	.61	.314	.37
Age	2888	21.97	5.80	.116	.14
Pell Grant	1386			-.046	-.05
First Generation	1058			-.050	-.05
Military Service	95			.086	.10
Part-time	645			-.117	-.14
American Indian	46			-.040	-.05
Black	142			-.050	-.06
Hispanic	720			-.073	-.09

*including only students who received a course grade

well. For every 186.7 minutes, or roughly 3 hours, of constructive dose, students on average attained 0.20 grade points for the course grade. Within this model, constructive dose included homework and discussions. Interactive dose was found to also have a significant positive effect on course grades. For every 495.1 minutes, or roughly 8 hours, of interactive dose, students on average attained 0.13 grade points for the course grade. Within this model, interactive dose included homework, recitations, and laboratory.

Considering assessments in the course sections such as quizzes, midterm exams, and final exams probably asked some recall and easier active level questions, the results for active dose were as expected. For example, active dose for homework refers to easier non-generative questions. These questions were graded with the same weight, in terms of points for the assignment, as more difficult problem-solving questions. Therefore, students spending time completing active level questions on homework were likely to improve course grades since course grade included active-level homework grades and assessment grades.

At the same time, homework assignments had a portion of constructive problem-solving questions that were likely to improve student learning. Indeed, constructive dose had a greater relative effect on course grades than active dose. Accounting for the ratio of dose time to improvement in grades, constructive dose likely improved course grades by an effect 112.4% greater, roughly double, than that of active dose. The relative effect of interactive dose was 75.5% less, about a quarter, than that of constructive dose and 47.1%, roughly half, less than that of active dose for improving course grades. A possible explanation for interactive dose having a reduced positive effect on final course grades, in contrast to final exams, could be due to the inability to ascertain whether students equitably co-constructed knowledge in both laboratory and recitation assignments since this study did not have any classroom observation level data. Considering the interactive dose of recitation assignments predicted a relatively equivalent effect as the constructive dose of homework on final exam scores, a likely explanation is that students may not have collaborated equitably and co-constructively on laboratory assignments. Further analyses of descriptives showed that the course sections averaged 335.58 laboratory interactive dose length and 30.64 recitation dose length. This discrepancy suggests that laboratory interactive dose likely had a much larger explanatory effect on course grade than recitation interactive dose.

Courses that were AOI or BSL were associated with greater course grades with students likely attaining 0.22 more grade points. Considering a large portion of course grade was calculated based on quizzes, midterms, and final exams, these results may be explained by prior research that shows that students perform better on online courses due to more accessible cheating than in F2F classes with proctored examinations (Bilen & Matros, 2021; Watson & Sottile, 2010).

Total SAT score within model 4 was associated with greater course grades with students likely attaining 0.23 more grade points for every 166.48 increase in total SAT score. In addition, incoming GPA was associated with higher course grades with students likely attaining 0.37 more grade points for every 0.61 increase in incoming GPA. These results were expected since students who performed well in prior academic courses were likely to also succeed in future academic courses.

In terms of student age, for every 5.8 years greater in age, students were likely to attain 0.14 more grade points. These results echo prior research that has shown that students older than 25 outperformed younger students in terms of course grades (Spitzer, 2000).

Pell grant recipient students were likely to attain 0.05 fewer grade points than those who did not receive a Pell grant. In addition, first generation college students were also likely to attain 0.05 fewer grade points than those whose family members previously attended college. Although the Pell grant is not the most accurate measure of student income since some low-income students do not receive them and some high-income students do receive them (Rosinger & Ford, 2019), in this study, the Pell grant acted as an approximate proxy for low-income status since I did not have access to detailed income level data for students. These results were expected since prior research has shown that student course grades were significantly predicted by income level (Brown & Burkhardt, 1999). In addition, the results for first generation students were expected since prior research has shown that first-generation students were likely to attain lower grade point averages for STEM courses (Whitcomb et al., 2021).

Students who identified as student service members or veterans on average attained 0.10 more grade points than those who did not identify as student service members or as veterans. These results contradict previous related research that found that students with military status performed more poorly in college with a lower GPA (Durdella & Kim, 2012). At the same time, the research did not specify differences for STEM GPA. There is not much literature that explores the relationship between military status and STEM course grades; these findings potentially contribute to the gap in literature regarding military status learners' success in STEM courses in college.

However, the literature does discuss effects of military service that may help explain our findings. In a qualitative inquiry study, Naphan and Elliot (2015) found that military service members, during their time in the military, developed a sense of collaborative task cohesion. Task cohesion refers to actions such as considering the needs of others ahead an individual’s needs and working diligently in order not to disappoint teammates. A possible explanation for why military service members were likely to attain more grade points could be, in part, due to this developed trait. While engaged in the interactive mode, military service members may be more effectively collaborating, possibly completing collaborative assignments with more care and diligence, which may explain the increase in attained grade points.

Table 13
Course Withdrawal

	Model 1 - e^{β}	Model 2 - e^{β}	Model 3 - e^{β}	Model 4 - e^{β}	Model 5 - e^{β}
Cox Snell R^2 / Nagelkerke R^2 / χ^2 sig.	.007/.017***	.007/.017***	.030/.079***	.136/.356***	.130/.302***
Active Dose	.999*	.999*	.999	1.000	1.001*
Constructive Dose	.999	.999	.999*	.997**	.998*
Interactive Dose	1.000*	1.000*	1.000	1.000	1.000
AOI or BSL		.946			
Total SAT			.998**	.998	
Incoming GPA			.474***	.634	.705*
Age				.939	.965*
Pell Grant				.916	1.022
First Generation				1.359	1.095
Female				1.714	1.795**
Military Service				.000	.065**
Part-time				33.311***	25.173***
Transfer				.638	1.044
2 or more Ethnicities				1.164	1.033
American Indian				2.647	1.844
AAPI				.800	.846
Black				1.228	2.024*
Hispanic				1.026	1.027

$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

Part-time students on average attained 0.14 points fewer grade points than full-time students. Given part-time students often have other responsibilities such as parenthood and careers, a plausible reason could be that part-time students might have other priorities in life that may make attaining more grade points more difficult. Tessema et al. (2014) found that an increase in working hours for students was correlated with a decrease in GPA.

As compared to students who identified as white, American Indian students on average attained 0.05 fewer grade points, Black students on average attained 0.06 fewer grade points, and Hispanic students on average attained 0.09 fewer grade points. These findings were in line with prior research that has established that students who identify as these three races/ethnicities earned lower grades in college (Perie & Moran, 2005).

Course Retention

Course retention was measured as a function of whether students withdrew from the course. Analysis of models for course retention excluded students who dropped the course within the first two weeks since students drop courses in the earlier weeks for several personal or logistical reasons. Data examining course retention are presented in table 13. Results are explained in terms of whether students withdrew from the course. Since the responses for course withdrawal was binary (nominal), as in students either withdrew or remained in the course for the full duration, a logistic regression was more appropriate for modeling the relationship between course withdrawal and the independent variables in this data. In multiple linear regression models, β, the standardized coefficient, is the commonly reported value for interpreting the relative effect of each independent variable on the dependent variable. In contrast, for multiple logistic regression models, e^{β}

or Euler's number to the power of the standardized coefficient, which represents the odds ratio for each predictor independent variable, is the commonly reported value for interpreting the relative effects within the model. Accordingly, the results are reported as odds ratios or likelihood for course retention. Additionally, the assumption for goodness of fit for all logistic models were met, resulting in insignificant values for the Hosmer-Lemeshow tests for each model; hence, none of the models were a poor fit.

Since course retention was potentially affected by all aspects of the course, I included all instructional activities for ICAP dose, except for lectures since passive dose had a VIF > 5. I included measures of course components, including homework, recitation, laboratory, and discussion posts.

Model 1. The first model's results indicated that student withdrawal from courses was significantly predicted by an overall model including ICAP dose [$R^2_{cox\&snell} = .007$, $R^2_{nagelkerke} = .017$, $p < .001$]. This model accounted for between 0.7% to 1.7% of the variance in student course withdrawal. The odds ratio coefficients, e^{β}, indicated that for each minute of active dose, students were 0.1% less likely to withdraw from the course (95% CI [.999, 1.000], $p < .05$). For each minute of interactive dose, students were marginally less likely to withdraw from the course (95% CI [.999, 1.000], $p < .05$). Constructive dose did not significantly predict the likelihood of student course withdrawal.

Model 2. The second model included mode of instruction, AOI or BSL, into the model. Results indicated that course withdrawal was significantly predicted by the updated model that included mode of instruction [$R^2_{cox\&snell} = .007$, $R^2_{nagelkerke} = .017$, $p < .001$]. This model accounted for between 0.7% to 1.7% of the variance in student course withdrawal. The mode of instruction - AOI or BSL did not significantly predict course withdrawal. Accordingly, I removed the mode of instruction from subsequent models.

Model 3. The third model included student incoming GPA. Results indicated that course withdrawal was significantly predicted by the updated model that included incoming GPA [$R^2_{cox\&snell} = .030$, $R^2_{nagelkerke} = .079$, $p < .001$]. This model accounted for between 3% to 7.9% of the variance in student course withdrawal. The odds ratio coefficients indicated that for each point value increase for total SAT, students were 0.2% (95% CI [.997, .999], $p < .001$) less likely to withdraw from the course and for each point value increase in incoming GPA, students were 53.6% (95% CI [.327, .688], $p < .001$) less likely to withdraw from the course. Controlling for mode of instruction, active dose and interactive dose did not have significant relationships with course withdrawal within this model.

Model 4. The fourth model included student demographics. Results indicated that course withdrawal was significantly predicted by the updated model that included student demographics [$R^2_{cox\&snell} = .136$, $R^2_{nagelkerke} = .356$, $p < .001$]. This model accounted for between 13.6% to 35.6% of the variance in student course withdrawal. For this model, the odds ratio coefficients indicated that part-time students were 3331.1% more likely to withdraw from the course (95% CI [17.587, 63.093], $p < .001$). Controlling for student demographics, incoming GPA and total SAT scores did not have significant relationships with course withdrawal within this model.

Model 5. The fifth model removed the total SAT score since the data only included total SAT scores for approximately half of the students. Results indicated that course withdrawal was significantly predicted by the updated model that included student demographics [$R^2_{cox\&snell} = .130$, $R^2_{nagelkerke} = .302$, $p < .001$]. This model accounted for between 13% to 30.2% of the variance in student course withdrawal. For this model, the odds ratio coefficients indicated that for each minute active dose, students were 0.1% more likely to withdraw from the course (95% CI [1.000, 1.001], $p < .05$) and for each minute of constructive dose, students were 0.2% less likely to withdraw from the course (95% CI [.997, 1.000], $p < .05$). Interactive dose did not have a significant relationship with course withdrawal in this model. For each point value increase in incoming GPA, students were 29.5% less likely to withdraw from the course (95% CI [.533, .933], $p < .05$). For each increase in age by 1 year, students were 3.5% less likely to withdraw from the course (95% CI [.935, .997], $p < .05$). In addition, female students were 79.5% more likely to withdraw from the course (95% CI [1.252, 2.571], $p < .01$), military students were 93.5% less likely to withdraw from the course (95% CI [.009, .485], $p < .01$), part-time students were 2417.3% more likely to withdraw from the course (95% CI [16.357, 38.741], $p < .001$), and race/ethnicity - Black students were 102.4% more likely to withdraw from the course as compared to students who identified as white (95% CI [1.137, 3.603], $p < .05$). Pell grants, first-generation status, transfer status, race/ethnicity - 2 or more, American Indian, AAPI, and Hispanic did not have significant relationships with course withdrawal.

Table 14
Summary for Course Retention

		N	Percent
Selected Cases	Included in Analysis	2670	84.9
	Missing Cases	474	15.1
	Course Withdrawal	218	8.2

Discussion of Findings for Course Retention. A statistical limitation of this model for course retention was that only a small percentage of students (8.2%) withdrew from the course, see table 14. Therefore, the results of these estimated likelihood ratios may underestimate the true effects and associations within the model for students who withdrew from the course. Accordingly, the following discussion of course retention is only an approximation of how different factors were likely to affect whether students withdrew from the course or not. Although this study cannot draw definitive conclusions, discussion of this model provides tentative implications that may warrant further study.

According to prior literature on college course retention, students withdraw for the following major reasons: a) low grades; b) too difficult to understand; c) not liking the design of the course; d) not liking the professor; and e) low interest in subject matter (Dunwoody & Frank, 1995). Therefore, in this discussion, I interpret the likelihood of effects based on these possible explanations.

In terms of the effect of ICAP dose, active dose predicted student withdrawal and constructive dose predicted student retention while interactive dose did not have a significant effect on course withdrawal. Considering interactive dose did not predict likelihood, the remaining activities that may have produced these effects were homework, lab, and discussion assignments. There are a few possible explanations for these effects: a) constructive dose may have deepened student understanding of chemistry in a manner that made the progression of the course easier to understand; b) students may have preferred the design of a course that had fewer basic homework and post-lab problems as well as more discussion assignments; and c) based on this study's findings on final exam score, students may have remained in the course since a decreasing active dose likely decreased exam scores whereas constructive dose likely increased exam scores, which could have had the effect of discouraging and encouraging students respectively.

Incoming GPA predicted student retention within the model. Prior research has found that the cumulative GPA of students prior to taking a college STEM course likely predicts student retention (Adams & Becker, 1990). Given incoming GPA and college GPA were related, these results are plausible. While institutional-level retention is a different characteristic than course-level retention, the relationship between these student outcomes may contribute to explaining these results. Possible explanations for these findings: a) students who have higher incoming GPA may be more likely to understand course material; and b) since higher incoming GPA was associated with higher course grades, students may have remained in the course due to satisfaction with course grades.

Student age also predicted student retention within the model. For every year increase in age, students were 3.5% less likely to withdraw from the course. Within this model, age also predicted higher course grades. Therefore, a possible explanation for the relationship between age and course retention may be due to older students' satisfaction with their course grades.

Female students were 79.5% more likely to withdraw from the course than male students. These results may be related to other findings in this study: a) females were likely to perform worse on final exams; and b) females were more likely to be part-time students, which was also associated with worse performance on final exams and course grade. A plausible explanation for these results is that the combination of these factors contributes to females withdrawing due to lower performance on assessments in the course.

Student service members and veterans were much less likely to withdraw from the course, 93.5% less likely than for non-military students. These results may be related to the findings in this study that showed military service students on average attained higher course grades. Since military service students attained higher course grades, this may have motivated them to remain in the course. However, the high reduction in withdrawal rates may not be fully explained by just higher course grades. There may be other reasons, such as motivational factors, for why military students may have been much less likely to withdraw than non-military service students.

Part-time students were more than 24 times more likely to withdraw from the course compared to full-time students. These results may be explained by the study's findings that part-time students perform more poorly on exams and attain lower GPA. At the same time, there may be other reasons for the very high likelihood of withdrawal. A plausible reason could be that part-time students might have other priorities in life that may make completing the course more difficult such as parenthood, careers, and other responsibilities of adult life (Benshoff & Lewis, 1992).

Students who identified as Black were about twice as likely to withdraw from the course compared to white students. These findings were consistent with a previous study that showed Black students withdrawing at a higher rate than white students (Cochran et al., 2014). While lower associated course GPA for Black students may explain these results, there were probably other reasons that may explain why students who identified as Black was the only race/ethnicity group to withdraw at a higher rate than white students considering Hispanic and American Indian students also attained lower course grades.

Table 15
Number of Weeks Participated

	Model 1 - β	Model 2 - β	Model 3 - β	Model 4 - β
R / Adjusted R^2 / ANOVA Significance	.426/.171***	.432/.172***	.399/.107*	.497/.195***
Active Dose	.256*	.239*	.069	.325*
Constructive Dose	-.142	-.129	-.274	-.137
Interactive Dose	.114	.097	.265	.129
AOI or BSL		-.078		
Total SAT			.021	
Incoming GPA			-.150	
Age				-.034
Pell Grant Recipient				-.041
First Generation				.120
Female				.017
Military Service				-.034
Part-time				.213**
Transfer				-.007
2 or more Ethnicities				.043
American Indian				.089
Asian American & Pacific Islander				.095
Black				.071
Hispanic				-.002

$^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$;

Course Attrition

Course attrition was measured as a function of the number of weeks students participated during the duration of the course and participated in the LMS. The number of weeks students participated were determined by counting the frequency of weeks that students logged into and performed mouse clicks in Canvas. Analysis of models for course grades included the full sample. Since attrition only applies for those who did not complete the course, the inclusion criteria were student withdrawal from the course. Analysis of models excluded students who dropped the course within the first two weeks since students drop courses in the earlier weeks for several personal or logistical reasons. Data examining course attrition are presented in table 15.

Since course retention was potentially affected by all aspects of the course, I included all instructional activities for ICAP dose, except for lectures since passive dose had a VIF > 5. I included measures of course components including homework, recitation, laboratory, and discussion posts.

Model 1. The first model's results indicated that the number of weeks students remained in the course were significantly predicted by an overall model including ICAP dose [$R = .426$, $R^2_{adj} = .171$, $p < .001$]. This model accounted for 17.1% of the variance in the number of weeks students remained in the course. The regression coefficients, β, indicated that active dose ($\beta = .256$, $p < .05$) had a significant positive relationship with the number of weeks students participated. Constructive and interactive doses had no significant relationships with the number of weeks students participated.

Model 2. The second model included mode of instruction, AOI or BSL, into the model. Results indicated that the number of weeks students remained in the course were significantly predicted by the updated model that included mode of instruction [R = .432, $R^2_{adj} = .172$, $p < .001$]. This model accounted for 17.2% of the variance in the number of weeks students remained in the course. The regression coefficient for AOI or BSL indicated that it did not have a significant relationship with the number of weeks students remained in the course ($\beta = -.078$, $p > .05$). Accordingly, I removed the mode of instruction - AOI or BSL from subsequent models.

Model 3. The third model included total SAT score and student incoming GPA. Results indicated that the number of weeks students remained in the course were significantly predicted by the updated model [R = .399, $R^2_{adj} = .107$, $p < .05$]. This model accounted for 10.7% of the variance in course grades. The regression coefficients for total SAT score ($\beta = .021$, $p > .05$) and incoming GPA ($\beta = -.150$, $p > .05$) indicated that they did not have a significant relationship with the number of weeks students remained in the course. Accordingly, I removed total SAT score and incoming GPA from subsequent models.

Model 4. The fourth model included student demographics. Results indicated that the number of weeks students remained in the course were significantly predicted by the updated model that included student demographics [R = .497, $R^2_{adj} = .195$, $p < .001$]. This model accounted for 19.5% of the variance in course grades. The regression coefficients for the final model indicated that active dose ($\beta = .325$, $p < .05$) and part-time status ($\beta = .213$, $p < .01$) had significant positive relationships with the number of weeks students remained in the course. Constructive dose, interactive dose, age, pell grant recipient, first- generation status, sex - female, military, transfer status, and race/ethnicity did not have significant relationships with the number of weeks students participated.

Table 16

Course Attrition - Descriptive Statistics - Significantly Related Variables

Model 4	N*	Mean	SD	β	β * SD_{DV}
# Weeks Participated	235	6.2000	3.11		
Active Dose	235	2432.5	523.3	.325	1.01
Part-time	180			.213	.66

*including only students who withdrew from the course

Discussion of Findings for Course Attrition. Active dose was found to have a significant positive effect on how long students remained in the course. For every 523.3 minutes of active dose, roughly 9 hours, students likely remained in the course for 1.01 more weeks in the course before withdrawing, see table 16. A possible explanation for this relationship could be that easier active-level homework and recitation problems may encourage students to stay in the course longer, which could be further reinforced by students likely attaining higher homework and recitation grades.

Part-time students likely remained in the course for 0.66 more weeks than that of full-time students. A couple plausible explanations for these results: a) part-time students may delay withdrawing from a course since they have a lower course load than that of full-time students; and b) part-time students might delay withdrawing from the course due to their busier personal life schedules. In general, some part-time students are parents or are working a job in addition to coursework. The difference of two-thirds of a week could possibly be attributed to fewer frequency in engaging in the course within a given week; part-time students may have withdrawn from the course only when time was available in their personal lives.

CHAPTER 6

Limitations

Trochim et al. (2016) states that generalizability, or external validity, of research is dependent on how representative the sample is compared to the general population. Due to the nature of the research sample, the generalizability of the findings of this study are limited to college students enrolled in college chemistry courses from a single large public university. The sample from this study does not account for variances in student population among different universities and subject matter. In addition, since the research design was a post-hoc data analysis using regression methods, the findings of this study do not establish causality, rather correlations, since there were threats to internal validity due to the non-experimental design of the study. The key criterion for determining causal inference is the experimentation of equivalent control and treatment groups created via random sampling and assignment a.k.a. randomized control trial (Ercikan and Roth, 2014). This study did not include a control group and did not use random sampling and assignment. Another limitation was that the findings for AOI and BSL course sections from Spring of 2020 to Spring of 2021 were based on instruction during the COVID-19 pandemic. Data collected during this period of time may have been affected by other factors such as increased stress and changes in access to university resources.

Within the course data, the results of this study were limited by the relatively small number of instructors who participated in the study. Of the 15 instructors who taught this introductory chemistry course, only 5 participated. All but one of the course instructors implemented nearly identical course sections, which resulted in effectively only six variations of ICAP dose in the study. As a result, the estimated effects of ICAP dose may have been skewed. Another limitation includes the level of data available for assign ICAP ratings for aspects of the course. Ideally, the data would have included video recordings of data and student work products thereby more accurately assessing how students were engaged in learning according to the ICAP framework. In addition, the data did not disclose how many times students watched and re-watched pre-recorded lectures for AOI students. This may have skewed the results for passive dose, which was based on the total length of pre-recorded videos, not the total time students watched and re-watched the videos. As previously mentioned, this study could not ascertain whether students reached the interactive mode of engagement since the finest level of evidence was at the course material level. Finally, this study did not ascertain the nature of final exam questions, which certainly influenced the relationship between ICAP dose and final exam grades. Final exams with more constructive level questions may have been more sensitive to generative dose and active level questions may have been more sensitive to non-generative dose.

Within student data, the findings and implications for race/ethnicity demographics were limited by the lack of distinction for international students, which may have affected results. In addition, a timely limitation for student data was the lack of gender identity options that list more than male or female. Future surveys should consider allowing for users to input their own gender as well as have options available other than the binary, male or female.

Table 17

Significant Findings for the ICAP Instructional Rubric

	Participation	Final Exam	Course Grade	Withdrawal	# Weeks Participated
Passive Dose	-	+			
Active Dose			+	+	+
excluding labs					
excluding labs/recitations					
Constructive Dose			+	-	
excluding labs	+				
excluding labs/discussions		+			
homework only	+				
discussions only	+				
Interactive Dose			+		
excluding labs		+			

CHAPTER 7

Implications and Recommendations

In this section, I summarize key findings from the study and discuss implications and recommendations for the field. I organized the discussion in the following order: a) the ICAP instructional rubric; b) course design; and c) learning analytics.

The ICAP Instructional Rubric

Overall, the results of this study affirm the value of the ICAP instructional rubric for predicting measures of student learning and performance. Although passive dose was found to be related to a reduction in student LMS participation and increased final examination grades, these results were explained by the inherent relationship between lectures and LMS participation as well as the relationship between final exam grades and the variance in student samples for those in AOI vs. BSL and F2F courses.

Results, see table 17 for an overview, showed that active dose was a significant predictor for increased course grades, increased course withdrawal, and greater number of weeks participated for students who withdrew. Increased course withdrawal may suggest that the courses that had more active-level problems in homework, laboratory, and recitation assignments were correlated with increased withdrawal rates.

In addition, results showed that constructive dose was a significant predictor for increased LMS participation, increased final exam grades, relatively larger increase in course grades, and decreased course withdrawal. These results affirm the ICAP theory, which predicts greater learning gains for constructive dose as compared to that of active dose. In addition, students were more likely to participate in the LMS for homework commensurate to the proportional dose of constructive level questions.

Finally, results showed that interactive dose was a significant predictor for increased final exam grades and a relatively smaller increase in course grades. The relatively smaller increase in course grades may have been due to the quality of collaboration in laboratory assignments. At the level of interaction between group members, this study could not ascertain the enacted quality of interactions. If collaboration was not equitable or partners did not co-construct knowledge, on average they would have received only fractions of constructive dose, which could explain why there was a relatively smaller effect than for that of constructive dose. For this same reason, for final exams, the positive effects of interactive dose were possibly underestimated. At a minimum, interactive dose had a marginally greater effect than constructive dose on final exam grades. These results add to the body of evidence confirming the ICAP theory's prediction for greater deeper learning.

These results provide evidence validating the use of ICAP instructional rubric for predicting student academic success, especially for the active dose, constructive dose, and interactive dose constructs. At the same time, further research is needed to validate the efficacy of passive dose through carefully designed studies that are more sensitive to variations in this construct. Additionally, further research should be conducted that ascertains the quality of interactive dose in order to elucidate the findings of this study.

Instructors may consider how much ICAP dose students are receiving for each mode of engagement. Based on the results of this study and discussion of findings, in my opinion, instructors do not need to alter their courses to be completely generative in order to achieve beneficial effects for students since the results showed that participation levels, final exam scores, and course grades were sensitive to relatively small variations in ICAP dose proportion. Certainly, passive and active modes of engagement have their place in college courses since instructors are concerned with students recalling information and concepts as well as developing and practicing skill sets, which are more closely related with non-generative modes of engagement. At the same time, instructors are likely concerned about students learning more deeply about the course content. In addition, course components other than lectures and laboratory experiments may warrant increased proportions of constructive and interactive dose. I would not recommend making all learning materials fully generative since students are also assessed on their ability to recall information and solve easier problems. Perhaps the ratio should be commensurate with how assessments are designed as well as the nature of the kinds of learning expected from the course. If a course has a heavy focus on recall, (e.g., a pharmacology course that requires the memorization of many drug names and interactions), instructors may consider engaging students in the lower two ICAP modes. Instructors who implement quizzes and examinations with more constructive questions may find that increasing the ratio of constructive and interactive dose has a beneficial effect on student performance. For instructors who intend to engage their students in more interactive dose, I recommend the implementation of expectations and accountability structures that facilitate and help ensure the equitable co-generation of learning among group members. This includes more monitoring of how students are collaborating.

While results for the relationship between ICAP dose and student participation are promising, further research needs to be conducted to elucidate explanations for these findings since there may be mediating or latent factors that may better explain the findings for levels of LMS participation. While I cannot make strong predictions on course retention based on the limited proportion of students withdrawing from courses, these results indicated that future research may be warranted for examining the relationship between ICAP dose and students withdrawing from courses. Constructive dose was associated with less withdrawal and active dose was associated with more withdrawal from the course.

In addition, considering the post-hoc data analysis design of this study, researchers may consider other experimental designs that provide evidence for causality such as randomized controlled trials or quasi-experimental pre-post nonequivalent groups design with propensity score matching in order to add to the body of evidence validating the use of the ICAP instructional rubric and further determining the relationship between ICAP dose and markers of student performance and outcomes.

At the same time, preliminary findings of a quasi-experimental pre-post non-equivalent groups design study suggest that the ICAP instructional rubric may be generalizable for secondary school instruction as well as provides evidence supporting the instrument's validity for predicting student learning. Echoing the results of this study, Ha et al. (2022) found that constructive dose had a relatively greater effect on student learning gains than active dose. The results of this dissertation along with the preliminary findings of Ha et al. suggest that further research implementing the ICAP instructional rubric may be warranted.

Finally, future research should continue to adapt the ICAP instructional rubric into more user-friendly iterations, such as Bloom's Taxonomy Pyramid and Webb's Depth of Knowledge Wheel, that are even more accessible for use by instructors. While figure 1 provides a helpful visual, the figure needs to be iterated in a manner that better guides how teachers should design their instruction.

Course Design and Evaluation

In this study, I examined the course design of multiple course sections of an introductory chemistry course at a large public university. Course components included lectures, homework, recitation, discussion, and laboratory. Even with relatively little variance in how the various versions of the course sections were implemented, student outcomes were likely significantly altered from variances in the ICAP ratings of the different components of the course. The findings of this study suggest that instructors should design their homework and recitation sections to contain a greater proportion of generative level questions. This relatively small change may have a great impact on the quality of deeper student learning and the level of student performance. In addition, based on the data in this study, instructors may want to consider analyzing the types of questions asked in online homework software; the two online software programs analyzed in this study asked mostly active-level questions whereas, in comparison, textbook assigned homework tended to have a greater proportion of constructive-level questions.

The findings of this study point to the predictive power of the ICAP framework for evaluating the design of college courses. Though more research is needed to further validate the use of the ICAP instructional rubric, the current evidence suggests that implementing small changes in ICAP dose can have great impact on student performance and outcomes. Universities may consider finding ways of implementing more rigorous course design and evaluation by integrating aspects of the ICAP instructional rubric into

Table 18

Significant Findings for Learning Analytics

	Participation	Final Exam	Course Grade	Withdrawal	Weeks Participated
AOI or BSL	+	+	+		
SAT Total	-	+	+	-	
Incoming GPA		+	+	-	
Age	-		+	-	
Pell Grant		-	-		
First Generation			-		
Female	+	-		+	
Military			+	-	
Part-time		-	-	+	+
Transfer					
Eth. - 2 or More					
Eth. - A. Indian		-	-		
Eth. - AAPI	+				
Eth. - Black			-	+	
Eth. - Hispanic		-	-		

currently used systems. For example, an institution using the Quality Matters program could replace “active learning” with aspects of the ICAP instructional rubric that describe how instructors can design courses that engage students more generatively and collaboratively. Moreover, administrators can utilize elements of the ICAP instructional rubric for assessing the quality of instruction and make subsequent evidence-based institutional decisions.

Learning Analytics

This study explored the relationships between ICAP dose and student characteristics, demographics, and mode of instruction. Table 18 summarizes significant findings for these relationships in this study. These results show that a number of student factors have positive and negative relationships with different markers of student success in college courses that may be of interest for instructors, administrators, and researchers. Some significant findings echoed the findings of prior research such as AAPI students performing better on final exams and female students performing more poorly. Other findings were contrary to prior research e.g., students with a military background attaining higher course grades. In addition, the results of this study also found significant relationships that are not well documented in the literature e.g., students with a military background were much less likely to withdraw from the course. Another interesting find that may warrant further study is to ascertain how older students, who were less likely to participate in the LMS, were more likely to attain higher course grades.

The findings of this study add to the body of learning analytics literature by showing how analysis of large data sets from educational institutions can elucidate trends within student populations as well as inform the design and evaluation of college courses.

For example, these significant findings for demographics can be utilized for designing improved learning experiences. There is a lot of LMS course data that is already available, education institutions only stand to be more informed about the administrative decisions they make by further commissioning the kind of research in this dissertation study.

REFERENCES

Adams, J.L., & Becker, W. E. (1990). Course withdrawals: A probit model and policy recommendations. *Research in Higher Education, 31*(5), 519-538. https://doi.org/10.1007/BF00992619

American College Testing. (2008). *2018 ACT/SAT concordance tables.* https://www.act.org/content/dam/act/unsecured/documents/ACT-SAT-Concordance-Tables.pdf

Ahmad, I., & Bokhari, M. U. (2013). The combine effect of synchronous and asynchronous e-learning on distance education. *International Journal of Computer Science Issues*, *10*(1), 546–550. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.696.5837&rep=rep1&type=pdf

Al-Samarraie, H. (2019). A scoping review of videoconferencing systems in higher education: Learning paradigms, opportunities, and challenges. *International Review of Research in Open and Distributed Learning*, *20*(3). https://doi.org/10.19173/irrodl.v20i4.4037

Alexander, B., Ashford-Rowe, K., Barajas-Murphy, N., Dobbin, G., Knott, J., McCormack, M., … Weber, N. (2019). *EDUCAUSE horizon report: 2019 higher education edition.* EDUCAUSE. https://library. educause.edu/-/media/files/library/2019/4/2019horizonreport.pdf?la=en&hash=C8E8D444AF372E705FA1BF9D4FF0DD4CC6F0FDD1

An, B. P. (2013). The influence of dual enrollment on academic performance and college readiness: Differences by socioeconomic status. *Research in Higher Education, 54*(4), 407-432. http://dx.doi.org/10.1007/s11162-012-9278-z

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, *89*(4), 369. https://doi.org/10.1037/0033-295X.89.4.369

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, *22*(3), 261-295. https://doi.org/10.1016/S0022-5371(83)90201-3

Anderson, J. R. (2005). *Cognitive psychology and its implications*. Macmillan.

Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives.* Longman.

Alibali, M. W., & DiRusso, A. A. (1999). The function of gesture in learning to count: More than keeping track. *Cognitive development, 14*(1), 37-56. https://doi.org/10.1016/s0885-2014(99)80017-3

Ausubel, D. P. (1977). The facilitation of meaningful verbal learning in the classroom. *Educational Psychologist, 12*, 162–178. https://doi.org/10.1080/00461527709529171

Ausubel, D. P. (1978). In defense of advance organizers: A reply to the critics. *Review of Educational Research, 48*, 251–257. https://doi.org/10.2307/1170083

Avcı, Ü., & Ergün, E. (2022). Online students' LMS activities and their effect on engagement, information literacy and academic performance. *Interactive Learning Environments, 30*(1), 71-84. https://doi.org/10.1080/10494820.2019.1636088

Baldwin, S., Ching, Y. H., & Hsu, Y. C. (2018). Online course design in higher education: A review of national and statewide evaluation instruments. *TechTrends, 62*(1), 46-57. http://dx.doi.org/10.1007/s11528-017-0215-z

Ballen, C. J., Salehi, S., & Cotner, S. (2017). Exams disadvantage women in introductory biology. *PLoS one, 12*(10), e0186419. https://doi.org/10.1371/journal.pone.0186419

Benshoff, J. M., & Lewis, H. A. (1992). *Nontraditional College Students.* ERIC Digest.

Bilen, E., & Matros, A. (2021). Online cheating amid COVID-19. *Journal of Economic Behavior & Organization, 182*, 196-211. https://doi.org/10.1016/j.jebo.2020.12.004

Bloom, B. S., Englehart, T., Furst, E., Hill, W., & Krathwohl, D. (1956). *A taxonomy of educational objectives, Handbook 1: Cognitive domain.* David McKay.

Bornschlegl, M., & Cashman, D. (2019). Considering the role of the distance student experience in student satisfaction and retention. *Open Learning: The Journal of Open, Distance and e-Learning, 34*(2), 139-155. https://doi.org/10.1080/02680513.2018.1509695

Bower, M. (2011). Synchronous collaboration competencies in web-conferencing environments – their impact on the learning process. *Distance Education, 32*(1), 63-83. https://doi.org/10.1080/01587919.2011.565502

Bower, M., Dalgarno, B., Kennedy, G., Lee, M., & Kenney, J. (2015). Design and implementation factors in blended synchronous learning environments: Outcomes from a cross-case analysis. *Computers and Education*, *86*, 1–17. https://doi.org/10.1016/j.compedu.2015.03.006

Braxton, J. M., Milem, J. F., & Sullivan, A. S. (2000). The influence of active learning on the college student departure process: Toward a revision of Tinto's theory. *The journal of higher education*, *71*(5), 569-590. https://www.jstor.org/stable/2649260

Brown, H. E., & Burkhardt, R. L. (1999). *Predicting student success: The relative impact of ethnicity, income, and parental education.* AIR 1999 Annual Forum Paper.

Card, D., & Payne, A. A. (2021). High school choices and the gender gap in STEM. *Economic Inquiry, 59*(1), 9-28. https://doi.org/10.1111/ecin.12934

Chen, W., & You, M. (2007, July). The differences between the influences of synchronous and asynchronous modes on collaborative learning project of industrial design. In Schuler D. (Ed.), *International Conference on Online Communities and Social Computing* (pp. 275-283). Springer. https://doi.org/10.1007/978-3-540-73257-0_31

Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science 1,* 73-105. https://doi.org/10.1111/j.1756-8765.2008.01005.x

Chi, M. T., Adams, J., Bogusch, E. B., Bruchok, C., Kang, S., Lancaster, M., ... Yaghmourian, D. L. (2018). Translating the ICAP theory of cognitive engagement into practice. *Cognitive science, 42*(6), 1777-1832. https://doi.org/10.1111/cogs.12626

Chi, M. T. H., Bassok, M., Lewis, M., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*, 145-182. https://doi.org/10.1207/s15516709cog1302_1

Chi, M. T. H., Deeleuw, N., Chiu, M. H., & Lavacher, C. (1994). Eliciting self-explanations improves understanding, *Cognitive Science, 18*(3), 439-477. https://doi.org/10.1016/0364-0213(94)90016-7

Chi, M. T. H., &Wylie, R. (2014). ICAP: A hypothesis of differentiated learning effectiveness for four modes of engagement activities. *Educational Psychologist*, 1-29. https://doi.org/10.1080/00461520.2014.965823

Clark, K. R. (2018). Learning theories: Constructivism. *Radiologic Technology, 90*(2), 180-182.

Cochran, J. D., Campbell, S. M., Baker, H. M., & Leeds, E. M. (2014). The role of student characteristics in predicting retention in online courses. *Research in Higher Education, 55*(1), 27-48. http://dx.doi.org/10.1007/s11162-013-9305-8

Cohen, M. A., & Greenberg, S. (2011). The struggle to succeed: Factors associated with the persistence of part-time adult students seeking a master's degree. *Continuing Higher Education Review*, *75*, 101-112. https://files.eric.ed.gov/fulltext/EJ967811.pdf

Conrad, N., & Patry, M. W. (2012). Conscientiousness and academic performance: A mediational analysis. *International Journal for the Scholarship of Teaching and Learning*, *6*(1). https://doi.org/10.20429/ijsotl.2012.060108

Cooper, R. P., & Shallice, T. (2006). Hierarchical schemas and goals in the control of sequential behavior. *Psychological Review, 113*(4), 887–916. https://doi.org/10.1037/0033-295X.113.4.887

Costa, P. T., Terracciano, A., & McCrae, R. R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality and Social Psychology, 81*(2), 322-331. https://doi.org/10.1037/0022-3514.81.2.322

Cronjé, J. (2006). Paradigms regained: Toward integrating objectivism and constructivism in instructional design and the learning sciences. *Educational Technology Research and Development, 54*(4), 387-416. https://doi.org/10.1007/s11423-006-9605-1

Crotty, M. (2020). *The foundations of social research: Meaning and perspective in the research process.* Routledge.

Darko, C. (2021). An evaluation of how students use Blackboard and the possible link to their grades. *SAGE Open*. https://doi.org/10.1177/21582440211067245

Dewey, J. (1923). *Democracy and education: An introduction to the philosophy of education.* Macmillan.

Dewey, J. (1986). Experience and education. *The Educational Forum, 50*(3), 241-252. https://doi.org/10.1080/00131728609335764

Dewey, J. (1933). Philosophy and civilization. *Philosophy, 8*(31).

Dietz-Uhler, B., & Hurn, J.E. (2013). Using learning analytics to predict (and improve) student success: A faculty perspective. *Journal of Interactive Online Learning, 12*(1), 17–26. http://www.ncolr.org/jiol/issues/pdf/12.1.2.pdf

D'Lima, G. M., Winsler, A., & Kitsantas, A. (2014). Ethnic and gender differences in first-year college students' goal orientation, self-efficacy, and extrinsic and intrinsic motivation. *The Journal of Educational Research, 107*(5), 341-356. https://doi.org/10.1080/00220671.2013.823366

Driscoll, M. (2000). *Psychology of Learning for Instruction*. Allyn & Bacon.

Durdella, N. R., & Kim, Y. K. (2012). Understanding patterns of college outcomes among student veterans. *Journal of Studies in Education, 2*(2). https://doi.org/10.5296/jse.v2i2.1469

Dunwoody, P.T. & Frank, M.L. (1995). Why students withdraw from classes. *The Journal of Psychology, 129*(5), 553-558. https://doi.org/10.1080/00223980.1995.9914927

Eddy, S. L., Brownell, S. E., & Wenderoth, M. P. (2014). Gender gaps in achievement and participation in multiple introductory biology classrooms. *CBE—Life Sciences Education, 13*(3), 478–492. https://doi.org/10.1187/cbe.13-10-0204

Ercikan, K., & Roth, W. M. (2014). Limits of generalizing in education research: Why criteria for research generalization should include population heterogeneity and uses of knowledge claims. *Teachers College Record, 116*(4), 1-28. https://doi.org/10.1177/016146811411600405

Ertmer, P. A., & Newby, T. J. (1993). Behaviorism, cognitivism, constructivism: Comparing critical features from an instructional design perspective. *Performance Improvement Quarterly, 6*(4), 50-72. https://doi.org/10.1111/j.1937-8327.1993.tb00605.x

Faw, H. W., & Waller, T. G. (1976). Mathemagenic behaviours and efficiency in learning from prose materials: Review, critique and recommendations. *Review of Educational Research, 46*, 691–720. https://doi.org/10.3102/00346543046004691

Feingold, A. (1994). Gender differences in personality: A meta-analysis. *Psychological Bulletin, 116*(3), 429-456. https://doi.org/10.1037/0033-2909.116.3.429

Field, A. (2013). *Discovering statistics using IBM SPSS statistics.* Sage.

Fischer, E. M. J. (2007). Settling into campus life: Differences by race/ethnicity in college involvement and outcomes. *The Journal of Higher Education, 78*(2), 125-161. https://doi.org/10.1080/00221546.2007.11780871

Fosnot, C. T. (2005). Constructivism revisited: Implications and reflections. *The Constructivist, 16*(1), 1-17.

Friedman, S. L., Klivington, K. A., & Peterson, R. W. (Eds.). (2013). *The brain, cognition, and education*. Academic Press.

Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends, 59*(1), 64–71. https://doi.org/10.1007/s11528-014-0822-x

Geary, D. C. (1995). Reflections of evolution and culture in children's cognition: Implications for mathematical development and instruction. *American Psychologist, 50*(1), 24-37. https://doi.org/10.1037/0003-066X.50.1.24

Gupta, P., & Cohen, N. J. (2002). Theoretical and computational analysis of skill learning, repetition priming, and procedural memory. *Psychological review*, *109*(2), 401-488. https://doi.org/10.1037/0033-295X.109.2.401

Ha, J., Boucher, N. S., Chi, M. T. H. (2022). The ICAP instructional rubric for evaluating lesson design and instructional implementation. Manuscript in preparation.

Henrie, C. R., Bodily, R., Larsen, R., & Graham, C. R. (2018). Exploring the potential of LMS log data as a proxy measure of student engagement. *Journal of Computing in Higher Education*, *30*(2), 344–362. https://doi.org/10.1007/s12528-017-9161-1

Henriques, L. (1997). *A study to define and verify a model of interactive-constructive elementary school science teaching.* The University of Iowa.

Hess, K. K., Jones, B. S., Carlock, D., & Walkup, J. R. (2009). Cognitive rigor: Blending the strengths of Bloom's Taxonomy and Webb's Depth of Knowledge to enhance classroom-level processes. https://files.eric.ed.gov/fulltext/ED517804.pdf

Hogan, K., Nastasi, B. K., & Pressley, M. (1999). Discourse patterns and collaborative scientific reasoning in peer and teacher-guided discussions. *Cognition and Instruction, 17*, 379–432. https://doi.org/10.1207/S1532690XCI1704_2

Hrastinski, S. (2010). How do e-learners participate in synchronous online discussions? Evolutionary and social psychological perspectives. In N. Kock (Ed.), *Evolutionary Psychology and Information Systems Research* (pp. 119–147). Springer. https://doi.org/10.1007/978-1-4419-6139-6_6

Igo, L. B., Bruning, R., & McCrudden, M. T. (2005). Exploring differences in students' copy-and-paste decision making and processing: A mixed-methods study. *Journal of Educational Psychology, 97*(1), 103-116. https://doi.org/10.1037/0022-0663.97.1.103

Jaggars, S., & Xu, D. (2016). How do online course design features influence student performance? *Computers and Education*, *95*, 270–284. https://doi.org/10.1016/j.compedu.2016.01.014

Johnson, L., Becker, S. A., Cummins, M., Estrada, V., Freeman, A., & Hall, C. (2016). *NMC horizon report: 2016 higher education edition* (pp. 1–50). The New Media Consortium. https://www.learntechlib.org/p/171478/

Jonassen, D. H. (1991). Objectivism versus constructivism: Do we need a new philosophical paradigm?. *Educational Technology Research and Development, 39*(3), 5-14. https://doi.org/10.1007/BF02296434

Katayama, A. D., Shambaugh, R. N., & Doctor, T. (2005). Computers in teaching: Promoting knowledge transfer with electronic note taking. *Teaching of Psychology, 32*, 129–131. https://doi.org/10.1207/s15328023top3202_9

Katona, G. (1940). *Organizing and memorizing: studies in the psychology of learning and teaching.* Columbia University Press.

Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction: Effects of direct instruction and discovery learning. *Psychological science, 15*(10), 661-667. https://doi.org/10.1111/j.0956-7976.2004.00737.x

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice, 41*(4), 212-218. https://doi.org/10.1207/s15430421tip4104_2

Kruger, D., Inman, S., Ding, Z., Kang, Y., Kuna, P., Liu, Y., … Wang, Y. (2015). Improving teacher effectiveness: Designing better assessment tools in learning management systems. *Future Internet*, *7*(4), 484–499. https://doi.org/10.3390/fi7040484

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation.* Cambridge University Press. https://doi.org/10.1017/CBO9780511815355

Mabrito, M. (2006). A study of synchronous versus asynchronous collaboration in an online business writing class. *American Journal of Distance Education, 20*(2), 93–107. https://doi.org/10.1207/s15389286ajde2002_4

Matz, R. L., Koester, B. P., Fiorini, S., Grom, G., Shepard, L., Stangor, C. G., ... McKay, T. A. (2017). Patterns of gendered performance differences in large introductory courses at five research universities. *AERA Open, 3*(4), 1–12. https://doi.org/10.1177/2332858417743754

Martin, F., Sun, T., & Westine, C. (2020). A systematic review of research on online teaching and learning from 2009 to 2018. *Computers and Education, 159*, 104009. https://doi.org/10.1016/j.compedu.2020.104009

Matthews, M. R. (1992). Constructivism and empiricism: An incomplete divorce. *Research in Science Education, 22*(1), 299-307. https://doi.org/10.1007/BF02356909

Mayer, R. E. (1996). Learners as information processors: Legacies and limitations of educational psychology's second.. *Educational Psychologist, 31*(3-4), 151-161. https://doi.org/10.1080/00461520.1996.9653263

Meece, J. L. (2002). *Child and adolescent development for educators.* McGraw-Hill Humanities, Social Sciences & World Languages.

Menekşe, M. (2012). *Interactive-constructive-active-passive: The relative effectiveness of differentiated activities on students' learning*. [Doctoral dissertation, Arizona State University]. https://keep.lib.asu.edu/items/151052

Menekse, M., Stump, G., Krause, S., & Chi, M. T. H. (2013). Differentiated overt learning activities for effective instruction in engineering classrooms. *Journal of Engineering Education, 102*, 346-374. https://doi.org/10.1002/jee.20021

Mestre, J. P. (2002). Probing adults' conceptual understanding and transfer of learning via problem posing. *Journal of Applied Developmental Psychology, 23*, 9–50. https://doi.org/10.1016/S0193-3973(01)00101-0

Meyer, K. A., Bruwelheide, J., & Poulin, R. (2009). Why they stayed: Near-perfect retention in an online certification program in library media. *Journal of Asynchronous Learning Networks, 13*(3), 129–145. https://doi.org/10.24059/olj.v10i4.1747

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81-97. https://doi.org/10.1037/h0043158

Morante, A., Djenidi, V., Clark, H., & West, S. (2017). Gender differences in online participation: Examining a history and a mathematics open foundation online course. *Australian Journal of Adult Learning, 57*(2), 266-293. https://files.eric.ed.gov/fulltext/EJ1148628.pdf

Morris, L.V., Finnegan, C.L., & Wu, S. (2005). Tracking student behavior, persistence, and achievement in online courses. *The Internet and Higher Education, 8*, 221–231. https://doi.org/10.1016/j.iheduc.2005.06.009

Naphan, D., & Elliot, M. (2015). Role exit from the military: Student veterans' perceptions of transitioning from the US military to higher education. *Qualitative Report*, *20*(2), 36-48. https://doi.org/10.46743/2160-3715/2015.2094

Neisser, U. (2014). *Cognitive psychology: Classic edition.* Psychology press. https://doi.org/10.4324/9781315736174

Norman, D. A., & Rumelhart, D. E. (1975). *Explorations in cognition.* W. H. Freeman & Company.

Piaget, J. (1952). *The Origins of Intelligence in Children.* W.W. Norton & Co. https://doi.org/10.1037/11494-000

Perie, M., & Moran, R. (2005). *NAEP 2004 trends in academic progress: Three decades of student performances (NCES 2005-464).* National Center for Education Statistics. https://files.eric.ed.gov/fulltext/ED485627.pdf

Pierrakeas, C., Xenos, M., Panagiotakopoulos, C., & Vergidis, D. (2004). A comparative study of dropout rates and causes for two different distance education courses. *The International Review of Research in Open and Distributed Learning, 5*(2). https://doi.org/10.19173/irrodl.v5i2.183

Powell, K. C., & Kalina, C. (2009). Cognitive and social constructivism: Developing tools for an effective classroom. *Education, 130*(2), 241-250.

Quality Matters (2018). *Specific review standards from the QM higher education rubric* (6th ed.). Quality Matters. https://www.qualitymatters.org/sites/default/files/PDFs/Standardsfrom theQMHigherEducationRubric.pdf

Reyes, J. (2015). The skinny on big data in education: Learning analytics simplified. *TechTrends*, *59*(2), 75–80. https://doi.org/10.1007/s11528-015-0842-1

Richardson, M., & Abraham, C. (2009). Conscientiousness and achievement motivation predict performance. *European Journal of Personality, 23*(7), 589-605. https://doi.org/10.1002/per.732

Rosinger, K. O., & Ford, K. S. (2019). Pell grant versus income data in postsecondary research. *Educational Researcher, 48*(5), 309-315. https://doi.org/10.3102/0013189X19852102

Schauble, L., Glaser, R., Duschl, R. A., Schulze, S., & John, J. (1995). Students' understanding of the objectives and procedures of experimentation in the science classroom. *The Journal of the Learning Sciences 4*(2), 131-166. https://doi.org/10.1207/s15327809jls0402_1

Schunk, D. H. (2012). *Learning theories, an educational perspective* (6th ed.). Pearson.

Schunk, D. H., & Meece, J. L. (Eds.). (1992). *Student Perceptions in the Classroom* (1st ed.). Routledge. https://doi.org/10.4324/9780203052532

Schwarz, B. B., Neuman, Y., & Biezuner, S. (2000). Two wrongs may make a right ... if they argue together! *Cognition and Instruction, 18***,** 461–494. https://doi.org/10.1207/S1532690XCI1804_2

Shattuck, J. C. (2016). A parallel controlled study of the effectiveness of a partially flipped organic chemistry course on student performance, perceptions, and course completion. *Journal of Chemical Education*, *93*(12), 1984-1992. https://doi.org/10.1021/acs.jchemed.6b00393

Shuell, T. J. (1986). Cognitive conceptions of learning. *Review of Educational Research, 56*(4), 411-436. https://doi.org/10.3102/00346543056004411

Skinner, B. F. (1965). *Science and human behavior.* Simon and Schuster.

Skinner, B. F. (1968). *The technology of teaching.* Appleton-Century-Crofts. https://doi.org/10.1177/002248716801900319

Skylar, A. (2009). A comparison of asynchronous online text-based lectures and synchronous interactive web conferencing lectures. *Issues in Teacher Education, 18*(2), 69-84. https://files.eric.ed.gov/fulltext/EJ858506.pdf

Smith, V. C., Lange, A., & Huston, D. R. (2012). Predictive modeling to forecast student outcomes and drive effective interventions in online community college courses. *Journal of Asynchronous Learning Networks, 16*(3), 51-61. https://files.eric.ed.gov/fulltext/EJ982673.pdf

Spitzer, T. M. (2000). Predictors of college success: A comparison of traditional and nontraditional age students. *Journal of Student Affairs Research and Practice, 38*(1), 99-115. https://doi.org/10.2202/1949-6605.1130

Stein, B. S., Littlefield, J., Bransford, J. D., & Persampieri, M. (1984). Elaboration and knowledge acquisition. *Memory & Cognition*, *12*(5), 522-529. https://doi.org/10.3758/BF03198315

Stowell, J. R., & Bennett, D. (2010). Effects of online testing on student exam performance and test anxiety. *Journal of Educational Computing Research, 42*(2), 161-171. https://doi.org/10.2190/EC.42.2.b

Stump, G. S., Li, N., Kang, S., Yaghmourian, D., Xu, D., Adams, J., ... Chi, M. T. (2017). Coding dosage of teachers' implementation of activities using ICAP: A video analysis. In Manalo, E., Uesaka, Y., & Chinn, C. A. (Eds.), *Promoting Spontaneous Use of Learning and Reasoning Strategies* (pp. 211-225). Routledge. https://doi.org/10.4324/9781315564029

Szeto, E. (2015). Community of inquiry as an instructional approach: What effects of teaching, social and cognitive presences are there in blended synchronous learning and teaching? *Computers & Education, 81*, 191–201. https://doi.org/10.1016/j.compedu.2014.10.015

Terry, W. S. (2015). *Learning and memory: Basic principles, processes, and procedures*. Psychology Press.

Tessema, M. T., Ready, K. J., & Astani, M. (2014). Does part-time job affect college students' satisfaction and academic performance (GPA)? The case of a mid-sized public university. *International Journal of Business Administration, 5*(2), 50. https://doi.org/10.5430/ijba.v5n2p50

Trafton, J. G., & Trickett, S. B. (2001). Note-taking for self-explanation and problem solving. *Human–Computer Interaction, 16*(1), 1–38. https://doi.org/10.1207/S15327051HCI1601_1

Trochim, W. M., Donnelly, J. P., & Arora, K. (2016). *The essential research methods knowledge base* (2nd ed.). Cengage.

Tulving, E. (1992). Memory systems and the brain. *Clinical Neuropharmacology*, *15*, 327A-328A. doi:10.1097/00002826-199201001-00169

Tulving, E. (1993). What is episodic memory? *Current Directions in Psychological Science*, *2*(3), 67-70. https://doi.org/10.1111/1467-8721.ep10770899

Van den Berg, M. N., & Hofman, W. H. A. (2005). Student success in university education: A multi-measurement study of the impact of student and faculty factors on study progress. *Higher education*, *50*(3), 413-446. https://www.jstor.org/stable/25068105

VanLehn, K., Graesser, A. C., Jackson, G. T., Jordan, P., Olney, A., & Rosé, C. P. (2007). When are tutorial dialogues more effective than reading? *Cognitive Science, 31*(1), 3–62. https://doi.org/10.1080/03640210709336984

Von Glasersfeld, E. (1984). An introduction to radical constructivism. In Watzlawick, P. (Ed.), *The invented reality* (pp. 17-40). Norton. http://vonglasersfeld.com/070.1

Vrasidas, C. (2000). Constructivism versus objectivism: Implications for interaction, course design, and evaluation in distance education. *International Journal of Educational Telecommunications, 6*(4), 339-362.

Vrasidas, C., & McIsaac, M. S. (1999). Factors influencing interaction in an online course. *American Journal of Distance Education, 13*(3), 22-36. https://doi.org/10.1080/08923649909527033

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Harvard University Press.

Vygotsky, L. S. (1962). *Thought and language.* MIT Press.

Wadsworth, B. J. (2004). *Piaget's theory of cognition and affective development.* Allyn & Bacon.

Wang, A. Y., & Newlin, M. H. (2002). Predictors of web-student performance: The role of self-efficacy and reasons for taking an on-line class. *Computers in Human Behavior, 18*, 151–163. https://doi.org/10.1016/s0747-5632(01)00042-5

Ward, M. E., Peters, G., & Shelley, K. (2010). Student and faculty perceptions of the quality of online learning experiences. *International Review of Research in Open and Distributed Learning, 11*(3), 57-77. https://doi.org/10.19173/irrodl.v11i3.867

Watson, G. R., & Sottile, J. (2010). Cheating in the digital age: Do students cheat more in online courses?. *Online Journal of Distance Learning Administration, 13*(1). https://www.westga.edu/~distance/ojdla/spring131/watson131.html

Webb, N. L. (1997). *Criteria for Alignment of Expectations and Assessments in Mathematics and Science Education. Research Monograph No. 6.* National Institute for Science Education. https://files.eric.ed.gov/fulltext/ED414305.pdf

Webb, N. L. (1999). *Alignment of Science and Mathematics Standards and Assessments in Four States. Research Monograph No. 18.* National Institute for Science Education. https://files.eric.ed.gov/fulltext/ED440852.pdf

Webb, N. M. (1989). Peer interaction, problem solving, and cognition: Multidisciplinary perspectives. *International Journal of Educational Research, 13*(1), 1–119.

Whitcomb, K. M., Cwik, S., & Singh, C. (2021). Not all disadvantages are equal: Racial/ethnic minority students have largest disadvantage among demographic groups in both STEM and non-STEM GPA. *AERA Open, 7.* https://doi.org/10.1177/23328584211059823

Wladis, C., Conway, K. M., & Hachey, A. C. (2015). The online STEM classroom—who succeeds? An exploration of the impact of ethnicity, gender, and non-traditional student characteristics in the community college context. *Community College Review, 43*(2), 142-164. https://doi.org/10.1177/0091552115571729

Wiggins, B. L., Eddy, S. L., Grunspan, D. Z., & Crowe, A. J. (2017). The ICAP active learning framework predicts the learning gains observed in intensely active classroom experiences. *AERA Open, 3*(2). https://doi.org/10.1177/2332858417708567

Xu, D., & Jaggars, S. (2011). *Online and hybrid course enrollment and performance in Washington State community and technical colleges.* Community College Research Center. https://doi.org/10.7916/D8862QJ6

Yamagata-Lynch, L. (2014). Blending online asynchronous and synchronous learning. *International Review of Research in Open and Distance Learning*, *15*(2), 189–212. https://doi.org/10.19173/irrodl.v15i2.1778

You, J. W. (2015). Examining the effect of academic procrastination on achievement using LMS data in e-learning. *Educational Technology & Society, 18*(3), 124–134. https://www.jstor.org/stable/jeductechsoci.18.3.64

You, J. W. (2016). Identifying significant indicators using LMS data to predict course achievement in online learning. *The Internet and Higher Education*, *29*, 23–30. https://doi.org/10.1016/j.iheduc.2015.11.003

Yuan, M., & Recker, M. (2019). Does audience matter? Comparing teachers' and non-teachers' application and perception of quality rubrics for evaluating open educational resources. *Educational Technology Research and Development, 67*(1), 39–61. https://doi.org/10.1007/s11423-018-9605-y

APPENDIX A

THE ICAP INSTRUCTIONAL RUBRIC

The ICAP Instructional Rubric

Table of Contents

[I. The ICAP Instructional Rubric]

[*Ia. Learning Activities*]

[*Ib. Learning Materials*]

[*Ic. Assessments*]

[II. Scoring Guidelines]

[*IIa. Considerations and FAQ*]

[III. Examples for Scoring Reference]

[*IIIa. Verb Phrases*]

[*IIIb. List of ICAP rated activities*]

The ICAP Instructional Rubric

	Passive	Active	Constructive	Interactive
Learning Activities	Activities are designed in a manner that involves students simply receiving information without any direction for manipulation.	Activities and teaching time are designed for all students to manipulate artifacts and/or other relevant learning materials without the generation of ideas.	Activities and teaching time are designed to foster generation and construction of ideas for all students.	Activities and teaching time are designed to foster equitable student cogeneration and co-construction of ideas.
Examples	Teacher lectures with a PowerPoint without requiring students to take notes. (This is still considered Passive even if some students take notes on their own accord) Class reads a text together with students participating by reading sections of the passage.	Students conduct a science lab in which the teacher gives directions and students follow the instructions one by one. Teacher lectures and asks students to take notes without any specification on how to take notes. Teacher asks students to work on math problems that have problems that mimic a known procedure.	A lecture requiring students to take Cornell notes, which includes summarization of lecture notes. Students are asked to work on a project that requires students to generate ideas. Students work on a worksheet that is constructive.	A classroom activity requiring students to collaboratively and equitably generate questions and predict answers. Students work on a group project that requires students to equitably generate ideas. A Socratic discussion in which all students are given the opportunity to prepare for the discussion and involves a structure that allows for equitable contribution.

Category	Passive	Active	Constructive	Interactive
Learning Materials	Learning materials are designed to be simply read or observed without any other overt action.	Learning materials are constructed to simply assess the retention of knowledge. These types of recall questions do not allow for the constructive generation of ideas and thoughts.	Learning materials allow for students to constructively generate responses, answers, or solve/answer generative, inference questions/problems.	Learning materials are structured for students to collaboratively and constructively generate responses, answers, or solve/answer generative questions/problems.
Examples	Powerpoint slides that only present information without asking students to manipulate anything. In-class reading materials that do not ask any questions.	A worksheet that asks students to answer without generation such as asking students to find answers in a textbook and copy down or asking students to simply identify or recall ideas and concepts. A math worksheet that asks students to answer problems based on a procedure the teacher has already enumerated. For example, skills-based or computational problems with no context.	A worksheet with questions that ask students to generate ideas such as summarizing (in their own words), explaining, reasoning, ranking, etc. A math worksheet that asks students to answer a novel problem or a word problem that requires students to use some inference for figuring out the solution. In particular, problems in context.	Everything in the column to the left with partners is Interactive as long as there are structures that expect equal contribution in constructing responses and products for all portions of the work products collaboratively.
Comments	Learning materials can be rated based on the type of question or activity as well as the rigor of the approach towards content as well. For example, “move”, “label”, and “measure” are Active verbs whereas “compare”, “rank”, and “explain” are Constructive verbs and “pair up” and “work together” are Interactive verb phrases. For example, multiple choice questions naturally fall within the domain of Active but can be Constructive if the question and prompt are difficult enough so that it requires students to **infer** in order to answer properly.			

Category	**Active**	**Constructive**	**Interactive**
Assessments	Student assessments are constructed to simply assess the retention of knowledge. These types of recall questions do not allow for the constructive generation of ideas and thoughts. These types of questions do not require students to infer.	Student assessments allow for students to constructively generate responses, answers, or solve/answer generative or inference questions/problems.	Student assessments allow for students to collaboratively co-generate responses, answers, or solve/answer generative or inference questions/problems.
Examples	Multiple choice questions on a test, true/false, fill in the blank with a word bank, matching, etc. Open response that only requires identification and answers an Active level question.	A test with open response questions asking students to summarize or explain a concept. A ranking task. Multiple choice questions with questions and answer choices that involve inferential thinking and reasoning.	A collaborative test or project that requires students to equitably co-generate.

Scoring Guidelines

Scoring is based on "ICAP Dose" (Stump et al., 2017). The procedure of scoring involves two steps: a) rate instructional activities and materials as engaging students in a particular ICAP mode; b) multiply the rating by the number of minutes students are engaged in those modes.

- The total number of minutes allocated for activities in a particular ICAP mode of engagement is termed **total ICAP dose**. The number of minutes should be rounded to the nearest whole number.
    - For example, if an activity has 64.8 minutes of an active activity and 35.2 minutes of a constructive activity, the total ICAP dose would be 65 active dose and 35 constructive dose.
- For worksheets, other classroom materials, and assessments, the total percentage of the number of items allocated in materials is termed material ICAP rating. The **material or assessment ICAP rating** converted to decimal form is then multiplied to the number of minutes spent on the worksheet or other classroom material in order to calculate the total ICAP dose.
    - If a worksheet has 3 active questions and 7 constructive questions, the material ICAP rating is 30% active and 70% constructive.
        - If this worksheet is used for 30 minutes in a given lesson, the dose is calculated by multiplying the number of minutes spent on the worksheet by the total dose percentage converted into decimal form. Here 30 minutes will be multiplied to 30% active, or 0.30, and 70% constructive, or 0.70, resulting in 9 active minutes and 21 constructive minutes total ICAP dose.
            - 0.30 active x 30 minutes = 9 active dose;
            - 0.70 constructive x 30 minutes = 21 constructive dose.
- Material or assessment ICAP rating cannot be rated as passive. Inherently, worksheets and assessments cannot be passive due to the active nature of completing them. If worksheets are implemented collaboratively, constructive items are rated as interactive. Active items remain active since interactive is defined as the co-generation of ideas and connections.
    - If a worksheet has 4 active questions and 4 constructive questions and 2 interactive questions, the material ICAP rating is 40% active, 40% constructive, and 20% interactive.
        - If this worksheet is used in a collaborative activity for 60 minutes in which **students are expected to equitably contribute to the co-generation of ideas**, the constructive portion of the total ICAP dose is changed to interactive. The total ICAP dose would be 24 active and 36 interactive.
            - 0.40 active x 60 minutes = 24 active dose.
            - (0.40 + 0.20) interactive x 60 minutes = 36 interactive dose.

Using the Best Available Evidence for ICAP dose ratings

There are varying grains of data that provide evidence for ICAP ratings within instruction. Finer grained data often elucidates coarser grained data. For example, a lesson plan may detail how much time teachers plan to spend on each activity. However, the actual enactment of the lesson may reveal that teachers spent different amounts of time on each activity. Moreover, the ICAP rating of implemented activities may differ from what might be assumed about the ICAP rating of the activity in the lesson plan. Therefore, finer grained data provide the best evidence for ICAP ratings of instruction and subsequently calculated doses. The non-exhaustive list below estimates the relative grain size of data; however, keep in mind that you, the rater determines grain size based on the data available.

In decreasing order of grain size:

- Course Description
- Syllabus
- Curriculum Map
- Course Schedule
- Lesson Plan
- Lesson/Course Materials
- Teacher Interview
- Student Interview
- Observation of Instruction
- Audio-recordings of Instruction
- Video-recordings of Instruction
- Student Work Products

For the most part, ICAP ratings should be based on multiple grains of evidence since finer grain sized evidence often lacks context. Ideally, ratings should be done in two phases, at the level of the teacher (intended) and then at the level of student (enacted). In phase 1, ratings and doses should be coded at the level of lesson/course materials and observations or recordings of instruction. In phase 2, student work should be scrutinized to see if the ratings from phase 1 are accurate.

Problem Solving Questions

Evaluating STEM-based problem solving/computation questions can be difficult to assess. Here are some guidelines.

1) **Problem solving** and **computation** problems are not equal (Fuchs et al., 2008). Problem solving involves the use of a language (here English) in order to construct a scenario that requires that students construct a problem model (internally or externally) in order to arrive at an answer e.g., word problems, problems in context. Assuming that the majority of students in a course do not have much prior knowledge or expertise about the topic, problem-solving questions in homework can be rated as Constructive.

2) **Procedural computations**, e.g., arithmetic (14 + 27), algebra (y = x +4; solve for x, convert 11 meters in nanometers), etc., should be rated as Active. Some examples of these questions ask students to convert, evaluate, estimate (sometimes worded as predict in STEM), calculate, solve, etc. without any other situational context.

3) Some problems do not fit neatly into these two aforementioned categories. Some problems may be just difficult enough to require some generation or inferential reason in order to solve a novel problem but might not be difficult enough to easily categorize as constructive-level. Usually, these types of problems are relatively easier to solve once students solve just one of them. Therefore, if there are multiple questions with similar questions structure, the first question can be considered constructive-level and all of the following highly similar questions can be considered active-level. If only one problem exists of this type, then err on the side of rating it constructive-level.

Other considerations

- Assessment ratings are not used for total ICAP dose but rather to determine the differential capacity of assessments to detect differences in deeper learning. For research purposes, assessment ratings can be used as exclusion/inclusion criteria.
- There are many activities that do not involve learning such as announcements, transitions, passing of materials, explaining procedures, etc. These activities should **not** be included in the total ICAP dose.
- If the time on an activity is undetermined, e.g., online courses, the rater should determine the time likely spent on activities, preferably ask the instructor of the course on what their expectations are for timing. In some cases, online homework software may suggest an estimated time for completion. These times can be used for determining the length of time students spend on homework.
- When rating material questions, make sure to look at reading materials associated with questions in order to determine if the questions can be directly answered based on information available in the reading material. If so, those questions should be rated as active even if they have wording that is constructive.
- Some questions and items have multiple parts, split each part into separate items.
- For student notetaking, base the rating on what the teacher directs and student work products. If a teacher does not ask students to take notes during a lecture, but some students do, that is still considered a passive activity. However, if the majority of students produced notes, then the rating should be elevated to active or constructive depending on the quality of notes.
- For lectures with PowerPoint slides, do not rate the lecture based on the slides. A lecture is a passive activity. If there is time for students to do work independently or collaboratively during a lecture, those times should be parsed out as different activities from the lecture and assigned separate ICAP doses.
- If students are asked to complete a task that is based on a PowerPoint, then use the rating of that slide.

FAQ

Q: When does an activity technically switch?

A: Whenever the instructor signals the shift.

Q: Do you use only the video for determining timing or do you also use the lesson plan?

A: Primarily use the video. However, if the timing is unclear, you can use the lesson plan to help determine the timing of how much time was spent on different activities.

Q: Should instructions for students to complete a task in a worksheet be considered an item?

A: Yes they should be as long as the directive likely (ideally observed) takes more than a minute to complete. If instructions have multiple parts in succession but are essentially a set of instructions, consolidate them into one item.

Q: What if timing data is unavailable?

A: Estimate the timing for activities; preferably ask the instructor for estimated times. Additionally, in some cases, standardized online materials have estimated times for completion that can be used.

Q: What if videos are unavailable?

A: Use the course materials i.e., course schedule, syllabus, homework assignments, etc., as well as ask instructors to clarify about timings in order to estimate the ICAP rating.

Q: What if only partial data is available, such as only 3 class sessions out of 12?

A: As long as there is evidence of consistent course structure i.e., course schedule or syllabus, rate the available material using the rubric with the assumption that the unknown class sessions, on average, will have similar ratings. For example, the average ICAP rating of the three homework can be applied to the remaining 9 weeks.

Q: What if questions in materials have multiple parts?

A: If the multiple parts just use seriation for the same task on multiple entities, this can be considered one question, e.g., calculate the slope for: a) $y = x$; b) $y = 2x + 3$, etc. If the different parts involve different questions or calculations, separate these into multiple questions, e.g., what is the symbol for sodium, how many neutrons does sodium contain, etc. If the multiple part is a directive such as explain? why or why not? justify with reasoning, etc., This can be considered as a modifier of the first part of the question and treated as a singular question, e.g., What is the force exerted on the block? Explain. This example should be rated as constructive because the modifier, explain, engages students in the constructive mode.

Examples for Scoring Reference

Examples of Verb Phrases

Passive	Active	Constructive	Interactive
	What is…?, Who is…?, Name…, Sketch…, Label…, Move…, Circle…, Use the…,	Why did…?, Explain your reasoning, Justify your position, What is the main idea of…? Suggest how…, Compare…, Determine…,	Debate with…, Discuss with…, Collaborate with…, Review with…, Talk with…, Act as a tutor or tutee, Interact with a computer tutor, Explain to another
If a teacher is posing phrases in direct instruction, oral or written, and only selecting a few students to answer, the teacher is engaging students in the Passive mode.	This includes any phrase that asks students to retrieve information or answers that are explicitly presented in the learning materials.	This assumes that whatever is asked is not explicitly presented in the learning materials. Students should generate answers that cannot be found in the learning materials.	These phrases indicate the Interactive mode if the activity or item is at the constructive level.

List of ICAP rated activities (Chi et al., 2022)

Passive No observable outputs	Reading, re-reading, observing, watching, looking, listening, tactile learning, turn-taking reading.
Active Observable outputs produced that are present in the original instruction or materials.	Clickers and polling, copying, enacting, mimicking, gallery walk, highlighting and underlining, identifying, generating a list, mapping, matching, choosing an option, measuring, lab activities, ranking, reading a selected passage or sentence aloud, reciting a poem, recall, rehearsing, review, total physical response (students moving their bodies as a response), guided notes with fill-in-the-blank, summarizing by only deleting irrelevant information, following step by step instructional, modeling by plotting or graphing, testing, re-testing, interleaved testing.
Constructive Observable outputs produced containing information beyond instruction or materials	Analogizing, comparing, contrasting, critiquing, criticizing, justifying, organizing, categorizing, concept-mapping, explaining to self, self-explanation, elaborating, explaining to another, posing questions, integrating two or more sets of information or ideas to form another set, integrating multiple media sources, predicting an outcome, taking notes in own words or paraphrasing, writing interpretive summary, writing interpretive review, connecting theories or concepts with everyday examples, designing experiments, hypothesizing explanations, modeling by aligning data with theories, deriving equations, reflecting, re-assessing, solving challenging problems with real world scenarios, project/problem-based learning, debugging, troubleshooting.
Interactive Partners reciprocally co-generating observable outputs that are Constructive	Asking and answering each other's questions, asking each other for justifications and explanations, debating with a partner by offering claims and justifications, improv theater or jazz ensemble, peer tutoring, reciprocal learning and teaching.

References

Chi, M. T. H., Boucher, N. S., & Ha, J. (forthcoming). The efficacy of learning strategies from the ICAP perspective. *International Encyclopedia of Education.* Elsevier Science.

Fuchs, L. S., Fuchs, D., Stuebing, K., Fletcher, J. M., Hamlett, C. L., & Lambert, W. (2008). Problem solving and computational skill: Are they shared or distinct aspects of mathematical cognition?. *Journal of Educational Psychology*, *100*(1), 30-47. https://doi.org/10.1037/0022-0663.100.1.30

Stump, G. S., Li, N., Kang, S., Yaghmourian, D., Xu, D., Adams, J., ... & Chi, M. T. (2017). Coding dosage of teachers' implementation of activities using ICAP: A video analysis. In Manalo, E., Uesaka, Y., & Chinn, C. A. (Eds.), *Promoting Spontaneous Use of Learning and Reasoning Strategies* (pp. 211-225). Routledge. https://doi.org/10.4324/9781315564029

APPENDIX B

INSTITUTIONAL REVIEW BOARD (IRB) APPROVAL LETTERS

EXEMPTION GRANTED

Margarita Pivovarova
Division of Educational Leadership and Innovation - Tempe
-
Margarita.Pivovarova@asu.edu Dear Margarita Pivovarova:

On 5/18/2021 the ASU IRB reviewed the following protocol:

Type of Review:	Initial Study
Title:	Evaluating postsecondary online learning and course design effects on student learning using the ICAP Instructional Rubric.
Investigator:	Margarita Pivovarova
IRB ID:	STUDY00012889
Funding:	None
Grant Title:	None
Grant ID:	None
Documents Reviewed:	• IRB Form_May2021_Ha_Pivovarova.docx, Category: IRB Protocol; • Jesse_Ha_Data_Request.pdf, Category: Other;

The IRB determined that the protocol is considered exempt pursuant to Federal Regulations 45CFR46 (1) Educational settings on 5/18/2021.

In conducting this protocol you are required to follow the requirements listed in the INVESTIGATOR MANUAL (HRP-103).

If any changes are made to the study, the IRB must be notified at research.integrity@asu.edu to determine if additional reviews/approvals are required. Changes may include but not limited to revisions to data collection, survey and/or interview questions, and vulnerable populations, etc.

Sincerely,

IRB Administrator

cc: Jesse Ha
Margarita Pivovarova
Jesse Ha

EXEMPTION GRANTED

Dear Brian Nelson:

On 10/15/2021 the ASU IRB reviewed the following protocol:

Type of Review:	Modification / Update
Title:	Evaluating postsecondary online learning and course design effects on student learning using the ICAP Instructional Rubric.
Investigator:	Brian Nelson
IRB ID:	STUDY00012889
Funding:	None
Grant Title:	None
Grant ID:	None
Documents Reviewed:	None

The IRB determined that the protocol is considered exempt pursuant to Federal Regulations 45CFR46 (1) Educational settings on 10/15/2021.

In conducting this protocol you are required to follow the requirements listed in the INVESTIGATOR MANUAL (HRP-103).

If any changes are made to the study, the IRB must be notified at research.integrity@asu.edu to determine if additional reviews/approvals are required. Changes may include but not limited to revisions to data collection, survey and/or interview questions, and vulnerable populations, etc.

REMINDER - All in-person interactions with human subjects require the completion of the ASU Daily Health Check by the ASU members prior to the interaction and the use of face coverings by researchers, research teams and research participants during the interaction. These requirements will minimize risk, protect health and support a safe research environment. These requirements apply both on- and off-campus.

The above change is effective as of July 29th 2021 until further notice and replaces all previously published guidance. Thank you for your continued commitment to ensuring a healthy and productive ASU community.

Sincerely,
IRB Administrator

cc: Jesse Ha
Margarita Pivovarova

APPROVAL: MODIFICATION

Margarita Pivovarova
Division of Educational Leadership and Innovation - Tempe
-
Margarita.Pivovarova@asu.edu

Dear Margarita Pivovarova:

On 1/25/2022 the ASU IRB reviewed the following protocol:

Type of Review:	Modification / Update
Title:	Evaluating postsecondary online learning and course design effects on student learning using the ICAP Instructional Rubric.
Investigator:	Margarita Pivovarova
IRB ID:	STUDY00012889
Funding:	None
Grant Title:	None
Grant ID:	None
Documents Reviewed:	None

The IRB approved the modification.

When consent is appropriate, you must use final, watermarked versions available under the "Documents" tab in ERA-IRB.

In conducting this protocol you are required to follow the requirements listed in the INVESTIGATOR MANUAL (HRP-103).

REMINDER - Effective January 12, 2022, in-person interactions with human subjects require adherence to all current policies for ASU faculty, staff, students, and visitors. Up-to-date information regarding ASU's COVID-19 Management Strategy can be found here. IRB approval is related to the research activity involving human subjects, all other protocols related to COVID- 19 management including face coverings, health checks, facility access, etc. are governed by current ASU policy.

Sincerely,
IRB Administrator

cc: Jesse Ha
Brian Nelson
Margarita Pivovarova
Jesse Ha