Rule Formation in Simulation-Based Discovery Learning: Optimized Clustering based on Levenshtein Edit Distance

by

Teeba Obaid

M.Ed., Athabasca University, 2016

B.A., Kwantlen Polytechnic University, 2009

Thesis Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

in the

Educational Psychology Program

Faculty of Education

SIMON FRASER UNIVERSITY

Spring 2024

Declaration of Committee

Name: **Teeba Obaid**

Degree: **Doctor of Philosophy**

Title: **Rule Formation in Simulation-Based Discovery Learning: Optimized Clustering based on Levenshtein Edit Distance**

Committee: **Chair: John Nesbit**
Professor, Education

Philip H. Winne
Supervisor
Professor, Education

Tenzin Doleck
Committee Member
Assistant Professor, Education

Engida Gebre
Examiner
Associate Professor, Education

Gregory Thomas
External Examiner
Professor, Faculty of Education
University of Alberta

Ethics Statement

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

a. human research ethics approval from the Simon Fraser University Office of Research Ethics

or

b. advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University

or has conducted the research

c. as a co-investigator, collaborator, or research assistant in a research project approved in advance.

A copy of the approval letter has been filed with the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Abstract

This study investigates how tools for formulating rules affect learning in a simulation of series electric circuits. Computer simulations can enhance exploratory learning but pose challenges to testing hypotheses, designing experiments, and interpreting data. I analyzed students' engineering tactics and search strategies in a simulation supplemented with tools guiding how to formulate rules. Participants were randomly assigned to a control or one of two experimental groups. Detectable strategy differences between groups were observed. Sequence analysis, leveraging Levenshtein edit distance, K-means clusters, silhouette coefficient, and generalized median method, revealed unique learning paths labeled Reinforced Confirmers, Dual-mode Strategy Diversifiers, Multi-strategy Jugglers, Self-regulated Revisers, and Methodical Integrators. This research contributes insights about effective instructional strategies for discovery learning in simulations, particularly how to improve knowledge integration and self-regulated learning in complex scientific domains.

Keywords: Simulation learning; discovery learning; knowledge integration; Levenshtein edit distance; K-means clustering; generalized median string; silhouette coefficient; electric circuits

Acknowledgements

I wish to express my profound gratitude to my supervisor, Dr. Phil Winne from Simon Fraser University. His unwavering guidance, invaluable feedback, and consistent encouragement have been pillars of support throughout the various stages of this dissertation—from research design to the challenging process of tool development. Particularly during the experimental phase of the tool, his patience and belief in my endeavors were unparalleled, serving as a continual source of motivation. His meticulous reviews and insightful editing further refined the culmination of this research journey. I am deeply appreciative of the foundation of trust and mentorship he has provided.

Special thanks to committee member, Dr. Tenzin Doleck from Simon Fraser University, whose invaluable insights into the research design and tool development enriched this study.

I am indebted to my colleague, Hosein Aghajani, for his meticulous work and collaboration in the realms of software development for machine learning algorithms and tool development. His proficiency has undeniably elevated the quality of this research.

I wish to express my heartfelt gratitude to Dr. Marcia Linn from the University of California, Berkeley, for her guidance during my time as a visiting researcher and her continued support through the Technology Enhanced Learning in Science Graduate Seminar. The invaluable critiques and insights from the members of Dr. Linn's research group have been crucial in refining the research plans, curriculum materials, assessments, results, and manuscripts associated with this study. Their collaborative spirit and the consistent, rigorous feedback they provided have been instrumental at every stage of this research journey.

On a personal note, I owe a debt of gratitude to my family, Hend and Zainab, for their endless encouragement and unwavering support throughout this academic endeavor.

Last but not least, I express my appreciation to all the participants who generously gave their time and insights to this study.

Table of Contents

List of Tables

List of Figures

Chapter 1. General Overview and Purpose of Study

1.1. Introduction

Simulation-based discovery learning uses computer simulations to promote exploratory learning and comprehension of intricate concepts like electric circuits and Ohm's law (Ma & Nickerson, 2006; Quellmalz et al., 2012). These systems' immersive virtual environment provides opportunities for learners to engage with scientific phenomena, design experiments, and test hypotheses. They have shown a potential to enhance conceptual understanding in domains like chemistry, allowing learners to visualize molecular structures and processes (Khan, 2011; Ma & Nickerson, 2006). However, despite appealing potential, learners using simulations for inquiry learning often encounter difficulties in testing hypothesis, designing experiments, interpreting data, and regulating their learning processes (de Jong et al., 1988; Friedler et al., 1990; Krajcik et al., 2000; Kuhn et al., 2000; Lewis et al., 1993; Njoo & de Jong, 1993; White, 1993, Bodemer, 2004). Bodemer (2004) found dynamic and interactive visualizations did not add much value over static ones, corroborating challenges that learners face in these areas. These challenges have been particularly noticeable in learning about series electric circuits, a topic notorious for its complexity and associated student misconceptions (Obaid et al., 2023a).

Insufficiently exploring alternative conditions and gathering limited evidence often leads to a poor understanding of the relationships between current, resistance, and voltage drop in series electric circuits. Occasionally, further misconceptions arise (Chambers et al., 1994; Obaid, 2023a). These difficulties not only depress motivation for learning but also hinder integrating domain-specific knowledge (Minstrell, 2000; Schauble et al., 1991). I agree with Bodemer's (2004) view: these difficulties are exacerbated in simulation-based learning environments when learners struggle to progress through stages of scientific inquiry. In this context, prior studies highlighted learners' comparative testing and inference-making are important to realize learning gains during simulation-based discovery learning (Obaid et al., 2023a). Bodemer's (2004) study supports this argument, as the effectiveness of active information integration was particularly evident for questions requiring transformation from graphical to textual representations.

Novice learners also encounter other challenges. Unless provided without appropriate guidance, they may rely on random search procedures or endure excessive cognitive load, leading to suboptimal learning (Kalyuga, 2011). Such challenges underscore the necessity for well-designed instructional methods in educational settings. In response to these challenges, there has been a growing adoption of computer-based learning environments, particularly interactive simulations. These simulations are designed to represent multi-step procedures and abstract concepts in a more engaging and understandable manner. The rationale behind this approach is to stimulate active learner engagement and encourage constructive learning processes, as suggested by de Jong & van Joolingen (1998), Rieber et al. (2004), and Schnotz et al. (1999). The aim is to cultivate deeper domain knowledge by enabling students to actively engage in scientific reasoning. This includes defining problems, formulating hypotheses, designing and executing experiments, and evaluating results. However, despite their potential, these simulations present their own set of challenges, as noted by Bodemer (2004), particularly in the stages of hypothesis formulation and data evaluation.

Moreover, the use of dynamic visualizations in these simulations, while intended to enhance learning, can sometimes contribute to the very challenges they seek to mitigate. Specifically, such visual elements can lead to cognitive overload or hinder self-regulated learning, as indicated by Lowe (1999) and Schnotz et al. (1999). Additionally, the integration of multiple types of external representations, such as text and images, often proves difficult for students. This difficulty can lead to disjointed knowledge structures, a concern echoed by Ainsworth (1999), Larkin & Simon (1987), and Mayer (1997, 2001). While there are supports designed to ease these challenges, their effectiveness has shown mixed results in empirical studies (van Joolingen & de Jong, 1991; Leutner, 1993; Njoo & de Jong, 1993; Swaak et al., 1998). A notable factor in these mixed outcomes is the learners' prior knowledge, or lack thereof, which is crucial for engaging effectively with complex visualizations, as discussed by Leutner (1993), Lowe (1999), and Schauble et al. (1991). This study investigates methods to reduce these challenges for learners using the WISE domain (https://wise.berkeley.edu/preview/unit/37434/node37) and the PhET electric circuit simulation (https://phet.colorado.edu/sims/html/circuit-construction-kit-dc/latest/circuit-construction-kit-dc_all.html) to investigate and learn about Ohm's law. I designed a bespoke rule formation tool to provide an interactive context for learners to explore electric circuits and test rules describing Ohm's law. This study delves into how learners' cognition and metacognition influence knowledge integration and self-regulated learning, and the potential to ameliorate misconceptions and foster effective hypothesis testing. The research underscores the critical role of scaffolding and guidance in rule formation, knowledge integration, and self-regulated learning, offering insights for future instructional design in similar educational settings.

Chapter 2. Literature Review

2.1. Overview

In chapter 2, I delve into several topics that provide essential context for addressing our research question. Firstly, I explore discovery learning and its applicability within simulation-based contexts, setting the stage for understanding how students engage with learning in this context. The integration of engineering approaches, aligned with the Next Generation Science Standards (NGSS), is another focal point, directly relevant to our research question as it investigates how students construct simulation circuits and plan investigations. Additionally, the discussion on search strategies in discovery learning underscores their importance in guiding learners during hypothesis generation and variable manipulation, critical aspects of problem-solving within simulation-based learning environments. I also examine heuristics, their benefits, and limitations, highlighting their relevance to the research question by exploring interventions that leverage heuristics to enhance learning within simulation-based contexts. Furthermore, this section addresses the selection of evidence for testing hypotheses, a core element of scientific discovery learning, and its connection to students' engineering and search strategies within simulation-based learning. I delve into the concept of rule formation in discovery learning, particularly in computer models and simulations, directly connecting to the research question by showcasing how students generate, apply, and modify rules to understand scientific phenomena. Self-regulated learning is introduced as an essential aspect that intertwines with rule formation in simulation-based discovery learning, highlighting the active role of students in their learning process. Finally, knowledge integration (KI) is discussed, emphasizing the merging of ideas for better comprehension, with direct relevance to the research question as KI explores how the rule formation tool aids students in making sense of observations and integrating new ideas. In essence, chapter 2 provides a comprehensive framework that connects these key topics to our research questions, offering insights into how students engage with simulation-based discovery learning and the factors and interventions influencing their learning paths and knowledge integration.

2.2. Discovery Learning and Simulation

Inquiry-based learning offers a robust framework for exploring the dynamics of how students form rules in simulation-based environments. Within this domain, discovery learning emerges as a particularly influential subset. Rooted in constructivist principles, discovery learning entrusts learners with the responsibility of navigating and interacting with learning materials to independently identify patterns and causal relationships, as highlighted by Reid et al. (2003). While inquiry-based learning can involve guided exploration based on questions, discovery learning particularly emphasizes positioning learners at the heart of the process, allowing them to uncover these patterns with minimal guidance, as emphasized by Alfieri et al. (2011). This educational approach reflects a gradual shift away from explicit instruction toward exploration and invention (Lazonder & Harmsen, 2016). In this study, I use inquiry learning to describe the environment and the instructional tasks to guide eliciting, distinguishing, and revising ideas through exploration tasks, prediction tasks, and investigation tasks; and I use discovery learning to refer to the cognitive process facilitated by the inquiry environment.

Notwithstanding the potential of discovery learning, the literature reveals an ongoing debate regarding its efficacy and limitations (Hmelo-Silver et al., 2007; Kirschner et al., 2006). Central to these discussions are discrepancies regarding the optimal structure and guidance for discovery tasks. Key aspects of guidance, such as timing, directiveness, and content, are debated, alongside the suitability of discovery learning across various learning contexts. Additionally, there is a discourse on balancing the demands of discovery learning with cognitive load considerations. The term 'discovery learning' itself remains broadly and inconsistently defined, primarily revolving around the notion of learners independently discovering target information or developing understanding from provided materials (Bruner et al., 1956).

Starting in the 1950s, research laid the groundwork for comparing discovery learning methods with other instructional forms, asserting the importance of self-guided comprehension (Bruner et al., 1956). More recent meta-analyses have elucidated conditions under which discovery learning, particularly in its assisted form, can outperform explicit instruction (Alfieri et al., 2011). While unassisted discovery learning has shown mixed results, enhanced discovery learning, featuring elements of scaffolding and guidance, has proven generally favorable compared to other instructional forms. This research deliberately focuses on discovery learning despite the ongoing debate and shift toward a framework labeled structured inquiry. The choice is informed by literature, including Alfieri et al. (2011), which illustrates that discovery learning, when effectively scaffolded, can foster more profound learning achievements. This nuanced understanding guides this study's approach, integrating elements of guidance within the discovery learning framework to balance autonomy with support, reflecting the mixed results in the efficacy of purely unassisted discovery learning.

While traditional explicit instruction might more efficiently promote acquiring facts, knowledge gleaned through discovery learning tends to endure longer and seep deeper into learners' values, cognitive skills, and self-constructed understanding (Dewey, 1910). This learning approach aligns well with the interactive, immersive nature of simulation-based learning, potentially enhancing the benefits of discovery learning. Moreover, the critical role of active engagement and learners' experiment design cannot be understated, especially in simulation-based discovery learning, where learners grapple with complex virtual environments (Chambers et al., 1994; de Jong & van Joolingen, 1998).

Discovery learning, while promising, can sometimes lead to learners feeling overwhelmed or failing to achieve the intended learning outcomes. This emphasizes the importance of carefully balancing discovery learning strategies. Reid et al. (2003) identified several difficulties learners encounter in scientific discovery learning (SDL), such as challenges in generating and adapting hypotheses, poorly designed experiments, difficulties in data interpretation, and problems with regulating the discovery learning process. These challenges highlight the need for sufficient support or scaffolding to facilitate effective discovery learning activities. SDL typically involves problem-solving activities and performing scientific experiments, comprising three main processes: representing problems and generating hypotheses, testing hypotheses through valid experiments, and reflectively abstracting and integrating discovery experiences (Reid et al., 2003).

Existing literature underscores the importance of worked examples and timely feedback in discovery learning environments, given inherent limitations of human working memory (Sweller et al., 1998). Furthermore, theory suggests effective discovery learning requires learners to engage constructively with the task and generate ideas beyond the given information (Kirschner et al., 2006). Reid et al. (2003) echo this, noting the quality of learner engagement with experimental activities significantly influences rule discovery and understanding.

Reid et al. (2003) proposed three types of learning support: interpretative support, experimental support, and reflective support. Interpretative support focuses on assisting learners with knowledge access and activation, hypothesis generation, and construction of coherent understandings. It plays a crucial role in activating relevant knowledge, enhancing problem representation, promoting hypothesis generation, and facilitating access to the knowledge base (Reid et al., 2003). Experimental support, on the other hand, aims to aid learners in the systematic and logical design of scientific experiments, prediction, observation, and conclusion drawing. While the overall effects of experimental support were less clear in Reid et al.'s (2003) study, it was found to have a positive effect on learners' intuitive understanding when supplemented by interpretative support. Experimental support alone did not significantly improve learners' experimental activities or overall learning outcomes, potentially due to the age of the participants (12-13 years old) (Reid et al., 2003). Nevertheless, the study underscored the importance of experimental activities that effectively guide learners to the correct rule, as these activities led to better outcomes in experiment design indices and the principle knowledge test.

Discovery learning coupled with simulation have been identified as potent teaching methods, with research indicating that discovery-based instruction enhances learning outcomes (Alfieri et al., 2011). Computer simulations have also been found to foster understanding of scientific concepts (Chambers et al., 1994; de Jong & van Joolingen, 1998). Yet, the critical role of active engagement and learners' experiment design cannot be understated, especially in simulation-based discovery learning, where learners grapple with complex virtual environments (Strand-Cary & Klahr, 2008).

Careful instructional design is required to ensure simulation and discovery tasks are well-structured, meaningful, and manageable. Learners need to be adequately supported throughout the process to ensure they do not become overwhelmed or confused. This involves offering appropriate scaffolding and feedback, as well as adjusting the level of guidance based on individual learner’s needs. The study by Correia et al. (2019) further supports the importance of scaffolding and support in simulation-based discovery learning. Their research focused on using the PhET simulation to teach gas laws on the submicroscopic level to secondary school students. The study incorporated a computer-assisted scaffolding system along with the simulation to facilitate conceptual understanding. The findings demonstrated that the combination of the scaffolding system and the simulation facilitated learners' conceptual comprehension of gas behavior on the submicroscopic level. Students reported positive learning experiences and highlighted the helpful features of the program, such as pop-up explanations, images, model explorations, guiding questions, diagrams, and feedback.

It is argued that with proper structure and guidance, the efficacy of discovery learning can be enhanced (Mayer, 2004; Alfieri et al., 2011). This involves optimizing the type and quality of support provided to learners, promoting better experiment design and rule discovery outcomes (Reid et al., 2003). By employing simulation-based discovery learning and utilizing a bespoke IF-THEN rule formation tool designed for this research, my study aims to enhance learning outcomes and promote the development of deep understanding of electric circuits. The tool aligns with Reid et al.'s (2003) findings regarding the importance of experimental support, providing structured guidance to facilitate the discovery learning process.

By leveraging the benefits of discovery learning and simulation-based environments, the intervention researched in this study provides learners with an interactive and immersive context intended to foster active engagement and the exploration of complex concepts. This approach aligns with constructivist principles, placing the learner at the center of the learning process. The study aims to contribute to the ongoing discourse around discovery learning, especially in a simulation-based environment, by exploring the role of experimental support and interpretative support in rule discovery. Through careful instructional design and the utilization of the IF-THEN rule formation tool, our research strives to shed light on the optimal conditions for successful discovery learning and provide actionable insights for educators and instructional designers in the field. Moving beyond the traditional approach of leaving learners to their own devices in simulation settings, this research underscores the imperative of structured support for deepening learners' understanding of relationships and enhancing their inferencing skills. This focus elevates the study's significance, advocating for a more guided learning experience within discovery and simulation-based education.

2.3. Engineering Approaches in Learning

In the context of the Next Generation Science Standards (NGSS) and the article by NGSS Lead States (2013), "engineering practices" refer to specific actions and skills students should engage to develop understanding of engineering principles and practices. These practices are an integral part of the NGSS framework, designed to integrate engineering into science education.

The NGSS identifies eight engineering practices students should participate in to enhance their knowledge and skills in engineering (NGSS Lead States, 2013). These practices serve as a foundation for students' engagement in the engineering design process and their ability to think and problem-solve like engineers. The eight engineering practices are (NGSS Lead States, 2013):

1. Defining problems: Students should be able to identify and define problems within the context of engineering design challenges. They should also be capable of establishing criteria for and constraints on potential solutions.
2. Developing and using models: Students should be proficient in creating and utilizing models to represent and simulate real-world systems or processes that are relevant to engineering design.
3. Planning and carrying out investigations: Students should engage in planning and conducting investigations to gather data and evidence for designing and improving engineering solutions.
4. Analyzing and interpreting data: Students should possess the ability to analyze and interpret data collected from investigations, identify patterns, and relationships, and make evidence-based decisions in engineering design.
5. Designing solutions: Students should demonstrate competence in developing and refining engineering designs using scientific and engineering knowledge, while considering factors such as criteria (specific requirements or standards that the design must meet), constraints, and trade-offs.
6. Engaging in argument from evidence: Students should construct arguments based on evidence and scientific reasoning to justify their engineering design decisions.

7. Constructing explanations and designing solutions: Students should be proficient in constructing explanations and designing solutions based on scientific and engineering principles.
8. Obtaining, evaluating, and communicating information: Students should be able to gather, evaluate, and communicate relevant information and data to support engineering design solutions. They should also engage in scientific and technical communication.

These engineering practices form the foundation for students' engagement in the engineering design process and their ability to develop a deeper understanding of engineering concepts and applications. By integrating these practices into science education, NGSS aims to foster critical thinking, creativity, and innovation among students (NGSS Lead States, 2013).

The application of engineering practices in this study was operationalized as "engineering approaches," which involved students' construction of simulation circuits and planning and carrying out investigations through comparative selections of variables. Additionally, the "searching strategies" were described as the process of forming rules during simulation learning. These engineering approaches reflect the NGSS engineering practices of "developing and using models" (practice 2) and "planning and carrying out investigations" (practice 3) using the electric circuit simulation and the rule-formation tool.

By incorporating these engineering approaches, this study provides students with opportunities to actively engage in the engineering design process and develop their understanding of engineering concepts and applications. The rule formation tool designed for this research and the PhET electric circuit simulation allow students to explore relationships between current, resistance, and voltage, analyze data, and make evidence-based decisions.

Furthermore, this study's focus on the impact of engineering approaches and search strategies on knowledge integration and self-regulated learning highlights the importance of understanding how students navigate the engineering design process and apply knowledge acquired in a simulated learning environment. This investigation contributes to the broader knowledge base on effective integration of engineering practices and simulation-based learning approaches, offering insights into instructional strategies that enhance students' critical thinking and problem-solving skills as engineers.

2.4. Search Strategies in Discovery Learning

Discovery learning is an instructional approach that emphasizes learners' active engagement in exploring and uncovering new knowledge. Central to the discovery learning process is the use of effective search strategies, which guide learners in manipulating variables and generating hypotheses to induce rules. The study by Bruner et al. (1956) serves as a foundational work in investigating learners' strategies for inducing rules through variable manipulation within a simulation-based context.

Bruner et al. (1956) identified and classified different search strategies employed by learners during discovery learning. These strategies include confirmation redundancy, simultaneous scanning, successive scanning, focus gambling, and conservative focusing. In confirmation redundancy, learners test alterations of the same instance repeatedly, seeking confirmation of a hypothesis. Simultaneous scanning involves considering all attributes simultaneously and eliminating or ruling out as many hypotheses as possible, optimizing for the most informative choice, while successive scanning entails attending to one attribute at a time. Conservative focusing involves changing only one attribute on each trial to investigate a single hypothesis, whereas focus gambling entails changing all but one attribute in each trial.

The study conducted by Farris and Revlin (1989) shed light on various search strategies employed by learners and their implications for the discovery learning process. By analyzing sequences of prediction, confidence measures, problem identification, rule formation, and knowledge integration, the researchers identified pathways through which learners built, revised, and integrated their understanding of concepts. These findings highlight the significance of search strategies in facilitating the formation and integration of rules during simulation-based discovery learning. The researchers found students consistently employed a disconfirmation strategy when assessing hypotheses in the rule discovery task. This strategy generates counterexamples that would be false if the hypothesis was true. However, when students were also required to generate hypotheses, they used a counterfactual inference strategy. This strategy assumes the hypothesized rule was false and generated examples consistent with an alternative hypothesis. While both strategies involve attempting to falsify a hypothesis, the disconfirmation strategy focuses on finding explicit contradictions to the hypothesis, while the counterfactual inference strategy involves assuming the falsity of the hypothesis and exploring alternative scenarios to assess its plausibility. The disconfirmation strategy directly targets the hypothesis, whereas the counterfactual inference strategy considers hypothetical conditions to evaluate the hypothesis indirectly. The results suggested that the selection of the hypothesis testing strategy depended on the logical requirements of the task and the desirability of the outcomes.

The study by Farris and Revlin (1989) contributes to understanding search strategies in discovery learning, emphasizing the importance of employing effective hypothesis testing strategies. It highlights the role of disconfirmation and counterfactual reasoning in facilitating the evaluation and generation of hypotheses during the rule discovery process. These strategies can enhance the formation and integration of rules during simulation-based discovery learning.

Klahr and Dunbar (1988) further expanded understanding of search strategies in discovery learning by proposing the dual search model of scientific discovery. Their studies involved subjects in a simulated scientific discovery context, aiming to uncover how a new function worked. They identified two main strategies for generating new hypotheses: searching the hypothesis space and searching the experiment space. Participants labeled Experimenters conducted experiments to test hypotheses and explored the experiment space, while Theorists searched the hypothesis space for an appropriate frame and proposed new hypotheses within the same frame. The Experimenters' goal was to discover the correct rule through systematic exploration of the experiment space. On the other hand, Theorists relied on prior knowledge and induction from outcomes to generate hypotheses. The results indicated Theorists tended to reach a solution faster than Experimenters. Their focus on searching the hypothesis space and leveraging prior knowledge allowed them to generate more plausible hypotheses, leading to more efficient discovery. In contrast, Experimenters, who relied on exploration of the experiment space, conducted a larger number of experiments to test hypotheses. The findings of Klahr and Dunbar (1988) highlight the importance of search in two problem spaces, the hypothesis space and experiment space, in scientific reasoning. Their proposed model offers a framework for understanding hypothesis formation and scientific discovery. The model emphasizes the interaction between learners' hypothesis generation, experimental design, and hypothesis evaluation. By incorporating the findings of Klahr and Dunbar (1988), this study recognizes the interaction between the hypothesis space and experiment space, as students navigate the complexities of variable manipulation and hypothesis generation. The simulation and the rule formation tool developed for this research provide a rich context for investigating learners' engineering approaches and search strategies. Students engage in constructing circuits and reading measurements to uncover relations involving current, resistance, and voltage. The rule formation tool allows students to specify hypotheses in the form of IF-THEN rules, facilitating their exploration of variables and predictions.

Together, the studies by Bruner et al. (1956), Farris and Revlin (1989), and Klahr and Dunbar (1988) provide valuable insights into the search strategies employed by learners during discovery learning. Effective search strategies enable learners to manipulate variables systematically, test hypotheses, and discover underlying rules and principles. By employing a range of search strategies, learners can explore different possibilities, consider multiple attributes, and engage in deep and meaningful learning.

In conclusion, investigations of search strategies in the context of discovery learning deepen our understanding of learners' approaches and their impact on knowledge integration and self-regulated learning. By building upon the existing literature, this study aims to improve our understanding of search strategies and their implications for effective instructional design and practice in discovery learning environments.

2.5. Heuristics and Discovery Learning

Heuristics play a significant role in discovery learning as evidenced by studies done by MacGregor and Cunningham (2008) and Strand-Cary and Klahr (2008) demonstrating the importance of heuristic searches in problem-solving situations within a discovery learning context. These heuristics, often considered as "rules of thumb," aid learners in dealing with complex problem-solving situations by providing cognitive tools to facilitate decision making. In recent literature, there's an understanding that teaching universal discovery learning techniques before embarking on the actual learning activity may have significant limitations. Scholars like Hodson (1998) argue that scientific methods aren't monolithic but rather vary depending on specifics of a science domain. This implies transfer from a general teaching scenario to a specialized context may not always happen seamlessly. Given the crucial role of heuristics in the learning process, integrating built-in support systems in learning environments becomes essential. Such support offers cognitive tools to learners, functioning to scaffold the learning process, whether by externalizing learning operations or structuring the task at hand (Lajoie & Derry, 1993; van Joolingen, 1999). Tools such as dedicated notebooks that aid structured note-taking, data manipulation tools, and explicit step-by-step forms are commonly integrated into learning environments to guide learners (Shute & Glaser, 1990; van Joolingen & de Jong, 1997).

Heuristics can enhance decision-making in discovery learning, particularly when exhaustive analysis of the problem or the context isn't possible due to incomplete information. However, it is essential to also teach the limitations of heuristics as they can sometimes lead to incorrect decisions. For instance, the heuristic VOTAT (vary one thing at a time), a principle related to experimental design, may not be effective in situations with interacting variables where it becomes necessary to change more than one variable simultaneously (Tsirgi, 1980; Zohar, 1995). Sanders et al. (2000) provided an inventory of heuristics suitable for simulation-based discovery learning environments. These heuristics, while specific enough to guide in specific instances, are general enough to be useful across multiple simulations. Some examples of these heuristics include: simplifying the problem, identifying and slightly modifying hypotheses, setting expectations, varying one thing at a time (VOTAT), and keeping track of what is being done.

The application of heuristics can be categorized into two main approaches: implicit and explicit. Implicit heuristics are subtly integrated into a learning environment, guiding learners through cues and guidance (van Joolingen & de Jong, 1997). These may result in successful behavior within the learning environment but can't be expected to translate to other domains or situations. On the contrary, explicit heuristics are explicitly and directly presented to learners, possibly fostering greater understanding and facilitating their application across different domains (Sanders, Bouwmeester, & Blanken, 2000).

A study explored the effectiveness of these two heuristic strategies within the context of a scientific discovery learning environment (Veermans et al., 2006). While both methods demonstrated significant gains in learners' domain knowledge, there were notable differences in learners' responses to these two approaches. In particular, the explicit heuristic condition proved more favorable for learners with lower initial domain knowledge. The study also revealed learners in the explicit heuristic condition tended to display more self-regulatory behavior, suggesting they incorporated the heuristics into their existing knowledge structures. By doing so, learners were better equipped to make decisions, evaluate their actions, and adjust their strategies autonomously. This ability to self-regulate and adapt reflects the depth with which the heuristics were internalized, highlighting their transformative potential when deeply embedded within the learners' cognitive framework.

In the current study, heuristics play a pivotal role as learners are guided to form rules within a simulation-based discovery learning setting. This environment harnesses a bespoke rule formation tool to aid students in crafting IF-THEN rules during the exploration phase of the study. The tool's usage, intertwined within the task of investigating current and voltage drop in series circuits, signifies a core aspect of the research. The students' iterative process of selecting the same value or changing values of variables, formulating and testing hypotheses within the simulation, and subsequently refining their rules, closely echoes heuristic principles within discovery learning (Sanders et al., 2000). The heuristics highlighted by Sanders et al. (2000) guide learners' decisions and actions within multifaceted, multi-step situations, akin to our simulation-based discovery learning environment. Analogous to the way heuristics assist learners in making informed decisions in the complex terrain of scientific discovery (Veermans et al., 2006), this study's IF-THEN rule formulation tool aims to support students in systematically orchestrating their experiments within the simulation. The current study seeks to understand how students' engineering approaches and search strategies in rule formation impact knowledge integration and self-regulated learning.

Research from Sanders et al. and Veermans et al. underlines the potential of explicitly integrating heuristics in the design of simulation-based learning environments, and emphasizes the value of guiding mechanisms within these simulations. Such insights, paired with burgeoning evidence on the utility of tools to assist students in systematically structuring discovery learning, prompted this study to conceptualize two distinct intervention approaches. The first intervention is grounded in the idea that a combination of tools might offer a more enriched experience. Thus, the decision table, facilitating the crafting of variable combinations and subsequent focused trials, is paired with rule boxes that automatically generate rules, aiming for a deeper, more comprehensive decision-making process. In contrast, the second intervention posits that perhaps a singular tool, in this case, the rule box, might suffice in guiding students efficiently, without the need for the added layer of the decision table. This distinction is critical. By having two separate interventions, the study can explore the incremental value, if any, brought about by the decision table. Is the combination of tools in the first intervention significantly more beneficial than just the rule box in the second? Or do students fare just as well with a more streamlined, singular tool? The ultimate aim is to discern which setup, be it multi-tooled or singular, best amplifies the effectiveness of heuristics within simulation environments.

2.6. Selecting Evidence to Test Hypotheses

Developing and testing hypotheses are crucial components of scientific discovery learning. The process of hypothesis testing involves forming hypotheses and gathering evidence to test and refine them (Klayman & Ha, 1987; Langley et al., 1987; Mahoney, 1976; Nisbett & Ross, 1980; Polya, 1954; Popper, 1959; Swann, 1984; Wason & Johnson-Laird, 1972). One commonly used task to study hypothesis testing strategies is the rule discovery paradigm introduced by Wason (1960). In this task, participants are presented with an initial set of items generated by a specific rule and are required to determine the underlying rule. Through multiple trials, participants propose new sets of items and receive feedback on whether their sets fit the rule, refining their hypotheses (Wason, 1960).

While many strategies are employed in hypothesis testing, one frequently observed approach is the positive test strategy. In this method, individuals tend to test instances they believe should align with the rule, often overlooking instances that contradict the hypothesized rule. This observation does not negate the diverse strategies identified by Bruner et al. (1956) but highlights one specific trend seen in certain learning contexts (Klayman & Ha, 1987). However, the positive test strategy has its limitations, such as confirmation bias and a narrow focus on instances that align with the hypothesized rule. Consequently, it often leads to hypotheses that fail to consider alternative possibilities (Wason, 1960). This tendency towards confirmation bias can result in individuals becoming overconfident in their incorrect rules (Wason, 1960).

Successful rule discovery involves considering alternative hypotheses and testing them (Klayman & Ha, 1987). By exploring different possibilities and testing alternative hypotheses, individuals enhance the process of hypothesis revision (Wason, 1960; Klayman & Ha, 1987). Notably, participants successful in rule discovery tasks demonstrate a greater tendency to direct their tests towards distinguishing explicit alternative hypotheses (Klayman & Ha, 1987).

In the current study, although the rule formation tool does not explicitly ask students to state their hypotheses, the process of making predictions and providing reasons indirectly aligns with hypothesis development and testing. Using the rule formation tool, students engage in a form of hypothesis testing by making predictions for each comparative trial. While a hypothesis offers a general explanation or assertion about potential relationships between variables, a prediction is a more specific statement about what will happen under particular conditions, based on that hypothesis. They formulate expectations about the relationship between the selected variables and the resulting current and voltage drop. The reasons provided for their predictions serve as justifications for their hypotheses or anticipated outcomes. The iterative nature of the tool allows students to refine their predictions and potentially modify their hypotheses based on previous comparative trials. The process of comparing their predictions with the actual measurements and analyzing their findings can be seen as a form of hypothesis testing and revision.

In the context of simulation-based discovery learning, incorporating insights from previous research on hypothesis testing provides a valuable framework for interpreting findings. The positive test strategy commonly observed in rule discovery tasks highlights the need to consider limitations of this strategy, including confirmation bias and a narrow focus on instances that fit the hypothesized rule (Wason, 1960; Klayman & Ha, 1987). Therefore, this study will explore how students' engineering approaches and search strategies relate to these patterns of hypothesis testing.

Furthermore, this study is designed to investigate the role of alternative hypotheses and tests of alternatives in the context of rule formation during simulation-based discovery learning. By examining whether students who successfully integrate knowledge and demonstrate self-regulated learning engage in more frequent and earlier testing of alternatives, insights can be gained into the importance of these factors for effective hypothesis testing and revision.

Moreover, the findings of this study can inform instructional design in simulation-based learning environments. By incorporating insights from previous research (Klayman & Ha, 1987), this study on rule formation in electric circuit learning can shed light on how students' engineering approaches and search strategies influence their hypothesis testing behaviors, knowledge integration, and self-regulated learning. These findings have potential to guide instructional practices that encourage students to adopt effective hypothesis testing approaches, thereby fostering a more comprehensive understanding of complex phenomena in simulation-based discovery learning environments.

In line with the importance of hypothesis testing, Brockbank and Walker's (2023) research on self-explanation provides additional insights into the learning process. Their findings suggest engaging in self-explanation enhances learning by directing attention and cognitive resources toward evidence that supports good explanations. This involves considering information that is broad, abstract, and consistent with prior knowledge, facilitating discovery and generalization. The rule formation tool in this study, while not directly seeking mechanistic explanations, engages students in creating comparative trials, making predictions, and rationalizing their choices. This encourages them to consider variable relationships and aligns with the overarching aim of fostering explanatory processes. Even without explicit prompts for mechanistic insights, the tool might lead students to make abstract references and think more generally.

While Brockbank and Walker’s (2023) experiments found explanation did not significantly impact hypothesis evaluation, it is important to note that in the tool used in the current study, evaluation remains an integral part of the iterative hypothesis testing process. In their study, participants in both the explanation and description conditions rated the target rule higher than other rules, indicating a general tendency to prioritize the correct rule regardless of whether they explained or described the evidence. However, in the context of the rule formation tool learners used in the current study, evaluation plays a crucial role. As students progress through different trials, they engage in an iterative process of hypothesis testing and evaluation. They develop hypotheses regarding the relationship between the constant and changing variables, make predictions based on those hypotheses, and subsequently compare their predictions with the actual findings of current and voltage drop.

2.7. Rule Formation in Discovery Learning

Computer models and interactive simulations can be potent tools in science education, offering students dynamic depictions of intricate scientific phenomena. They let students manipulate, explore, and observe system behaviors, enhancing their understanding. Central to harnessing these models effectively is the concept of rule formation. Klahr and Nigam (2004) and Wiese and Linn (2021) emphasized that rule formation, whether through direct instruction or discovery learning, is pivotal in early science instruction. It is vital for learners' cognitive development and grasping scientific concepts.

Rule formation in the context of computer models and interactive simulations plays a critical role in facilitating students' engagement with scientific concepts and phenomena. The rules operationalized by a model define the relationships, interactions, and constraints that dictate how the system or phenomenon behaves. By comprehending and manipulating these rules, students can develop a deeper understanding of the scientific principles and mechanisms at play.

Wiese and Linn (2021) explored the importance of understanding rule formation in students' interactions with computer-based simulations of scientific models. The study aimed to decipher how students perceive and utilize the foundational rules that drive these computer simulations to bolster their understanding of scientific phenomena. To be clear, by "rules," the study refers to the fundamental principles or behaviors that a simulation or model adheres to, while "model" refers to the entire simulation itself. The model is essentially the manifestation of multiple rules. Wiese and Linn devised a computational modeling inventory to assess students' computational thinking skills, specifically in rule formation. One key component of this inventory was the rule sorting task. Here, students were presented with a set of proposed rules and asked to sort them into two categories: rules used by the computer and rules not used by the computer. In essence, students were not directly programming; instead, during the rule sorting task presented by Wiese and Linn, students were tasked with deducing the operational rules of the computer model. This deduction was based on the patterns, outcomes, and changes they could directly observe or interact with when running the simulation. Presented with a set of rules written in plain language, which highlighted key model behaviors and common misconceptions, students sorted rules into two categories: rules used by the computer and rules not used by the computer. These rules, crafted to be understandable by middle school students, were centered on the object behaviors of the model and were devoid of specific code references. Thus, rather than engaging with or manipulating actual code, students interpreted how the model behaved under certain conditions, attempting to discern the foundational rules that governed these evident behaviors. The researchers used two models, the plant growth model and the global climate model, for the Rule Sorting task. For the plant growth model, students were instructed to run the model for 4 weeks with the light on and then turn the light off. This allowed them to observe the behaviors targeted by the rule sorting questions, such as the increase in total glucose made when the light is on and the decrease in total glucose stored when the light is off. Similarly, for the global climate model, students were instructed to run the model first without greenhouse gasses and then press the "Run Factory" button to observe the behaviors related to greenhouse gasses.

The results of the rule sorting task showed students had varying levels of success in identifying the model rules. Students were most successful in identifying rules that manifested in clear, visible behaviors when the model ran, deemed “directly observable in the model,” or those that the model clearly contradicted. For example, rules like "When infrared radiation hits a greenhouse gas, it will change direction" in the global climate model or "Total glucose used always increases" in the plant growth model were more likely to be sorted correctly. However, students struggled with sorting rules that referred to variables not included in the model or understanding the distinction between a scientific concept and its operationalization in the model.

The study found some students thought a model could follow contradictory rules, indicating a lack of understanding of the consistency required in a model's behavior. Students also had difficulty distinguishing between a science concept and the way it is operationalized in a model. They often believed models included more variables than they actually did, leading to misconceptions about the accuracy and representation of models. Furthermore, the study revealed students had difficulty recognizing emergent patterns in models. Emergent patterns refer to complex outcomes or behaviors that arise from the combined interactions of the model's rules, often leading to phenomena that aren't explicitly outlined by any single rule. They often accepted redundant rules that were not necessarily part of the model but were consistent with the observed behaviors. This lack of distinction between necessary rules and emergent patterns hinders students' understanding of how complex system behaviors can arise from simple underlying rules. Overall, the rule sorting task provided insights into students' understanding of computer models and their ability to recognize the underlying rules and behaviors. The findings highlight the importance of helping students connect observable behaviors in models with the underlying rules and distinguish between scientific concepts and their operationalization in models. The results of the Rule Sorting task indicated that students could thoughtfully engage with the concept of model rules but also revealed difficulties with decomposition, algorithms, and abstraction in models.

Wiese and Linn also incorporated new rule questions, which prompted students to consider the effects of implementing new, hypothetical rules in the model. "New rule" questions asked students to imagine how a modified (and scientifically incorrect) computer model would behave if certain rules were altered or added. It is important to note that students were not physically altering the actual code but were rather speculating on the effects of hypothetical changes to the rules governing the model. These questions aimed to prompt students to engage with incorrect ideas and assess their understanding of the science content.

The findings suggested allowing students to change the rules in a model could help them see the connections between model rules and behaviors, thus improving their computational thinking and understanding of the underlying science. Students were asked to make predictions about how the new rule would change the model's behavior and then observe the actual effects of the new rule. This approach aimed to deepen students' science understanding and encourage meaningful engagement with the modified model. By comparing students' responses to typical questions (related to correct models) and new rule questions (related to modified models), the researchers investigated whether different question types elicited different kinds of ideas. The typical questions focused on explaining relationships between variables portrayed in the models, while the new rule questions asked students to reason through how the modified model would behave. For example, in the global climate model, a typical question asked students to explain the relationship between greenhouse gasses and temperature, while a New Rule question asked what would happen if solar radiation always bounced off Earth's surface. Students could reason that, in the modified model where the new rule dictates that all solar radiation is reflected, the temperature would be much colder than in the original model. Similarly, in the chemical reactions model, a typical question asked students to explain the relationship between molecular movement and temperature, while a New Rule question asked how the formation of water would be affected if single hydrogens bonded with each other whenever they collided. The new rule questions delved into the mechanisms behind the phenomena in the models and challenged students to draw conclusions from animations. Correct responses to these questions indicated that students noticed crucial details and understood why certain phenomena occurred in the models.

Building upon the work of Wiese and Linn, the present study aims to further explore the concept of rule formation, specifically in the context of simulation-based discovery learning. This study focuses on the construction of series electric circuits using a physics simulation and an if-then rule formation tool. Students will interact with a physics simulation and an if-then rule formation tool to generate, apply, and modify rules that help them understand scientific phenomena in this domain. Simulation-based discovery learning serves as a platform for students to interact with dynamic and interactive models of series electric circuits, gaining hands-on experience and developing an intuitive understanding of the underlying scientific principles. The if-then rule formation tool complements the simulation by allowing students to formulate rules based on their observations and reasoning, further deepening their understanding and engagement with the topic. Similar to Wiese and Linn's New Rule questions, the if-then rule formation tool in this study prompts students to imagine and explore modified scenarios within the electric circuit simulation. By modifying the circuit's components or parameters, students can investigate the effects of changing the underlying rules or conditions on the circuit's behavior. This approach not only encourages students to engage with incorrect or alternative ideas but also provides an opportunity for them to deepen their understanding of the relationships involving current, resistance, and voltage.

2.8. Self-regulated Learning

The importance of self-regulated learning (SRL) in shaping learners' abilities to take control of their learning process has been widely acknowledged in educational research. This study, exploring rule formation in simulation-based discovery learning is firmly rooted in the concepts of SRL.

In SRL, learners constantly make decisions about their learning pathways, manage their attention, select learning strategies, and review their understanding (Winne, 2022). They are seen as agentic, meaning that they are actively involved in their learning process. This engagement and ownership over the learning process echoes Bandura's (1997) concept of agency, part of the theoretical grounding of SRL. Winne (2022) proposes self-regulated learning can be seen as learners doing learning science, a perspective that builds upon the works of Bandura (1997) and Ericsson and Harwell (2019) among others. Essentially, learners are positioned as scientists conducting experiments on their learning processes, observing, monitoring, and adjusting their cognitive operations to optimize their learning outcomes.

Simulation-based discovery learning, the focus of this study, aligns with the concept of SRL by positioning learners as active agents who make decisions and construct knowledge in a simulated environment. Moreover, the IF-THEN rule tool integrates with Winne's (2023) proposition of learners doing learning science, enhancing their ability to apply the scientific method in their learning process.

Winne (2022) emphasizes that developing productive SRL requires deliberate practice, which can be supported through learning analytics based on data generated by learners with strong correspondence to theoretical constructs, i.e., tracing those constructs. Trace data and learning analytics offer a means for learners to track their learning process, enabling self-regulation and fostering a deeper understanding of how they learn.

The application of a physics simulation and an IF-THEN rule tool in this study can potentially facilitate deliberate practice thus strengthening the learners’ SRL capabilities. Furthermore, these tools may help address the challenges learners often face, such as misconceptions, lack of data, and inadequate analytical methods (Winne, 2022).

Winne (2022) proposes a framework consisting of three models: the SMART model (searching, monitoring, assembling, rehearsing, translating) of cognitive operations, the COPES model (conditions, operations, product, evaluations, standards) of learning tasks, and the AEIOU (attributions, efficacy expectations, incentives, outcome expectations, utility) model of reasoning about learning as a motivated event. These models provide a structure for understanding and measuring SRL within a simulated environment like that used in this study.

The study's IF-THEN rule tool operationalizes the link between operations and products of the COPES model and also aligns with the SMART operations model. It allows learners to manipulate information and execute operations to produce desired learning outcomes. This could lead to the enhancement of learners' self-regulation skills and their ability to understand and manipulate conditions within their learning environment. The simulation augmented by the IF-THEN tool for proposing rules may also provide an environment for learners to practice and inferencing operations. For example, learners engage in searching when they explore the simulated environment and identify relevant information. They rehearse when they repeat experiments or scenarios to confirm the validity of their if-then rules. They translate when they convert their observations into IF-THEN rules. In essence, the simulation-based discovery learning process exemplifies the self-regulated learning as learners doing a learning science model.

Furthermore, learners' engagement with the rule tool and physics simulation potentially aligns with the AEIOU model. By interacting with these tools, learners could possibly heighten their awareness of motivations related to their learning, thereby expanding topics of their metacognition.

Moreover, this study's design, which incorporates a physics simulation and an IF-THEN rule tool, is instrumental in gathering trace data about learners' SRL behaviors. This design aligns with Winne's (2023) emphasis on the importance of data in understanding and promoting SRL. The generated trace data can help identify patterns, assess the effectiveness of SRL strategies, and support the development of tailored learning analytics.

The challenges Winne (2022) highlights, particularly in relation to learners' lack of expertise in applying the scientific method, erratic records of their learning, and the shortcomings of the data available for guiding metacognition, could be potentially addressed through the use of technology-enabled learning environments like the physics simulation tool and the IF-THEN rule formation tool.

The IF-THEN rule formation tool used in this study is a direct application of the outcome expectation concept discussed in Winne's paper. As learners interact with the physics simulation, they form and refine if-then rules based on their observations and experiences. These if-then rules essentially form the "theories" that the learners, as "scientists", use to understand and predict the behavior of the simulated system. The practice of creating these rules, testing them, and refining them can be seen as a form of deliberate practice, which Ericsson & Harwell (2019) argue is necessary for the development of expertise.

The study's use of rule formation logs could help capture comprehensive data about learners' actions and decisions in the simulated environment, which would address the problem of "erratic records about how they learn." This data could provide insights into the SMART operations enacted by learners, the conditions under which they occur, and the efficacy of these operations in producing the desired learning outcomes. The data could also be used to create future personalized feedback and recommendations to guide learners in their SRL processes.

2.9. Knowledge Integration

This study is primarily anchored in the framework of Knowledge Integration (KI), a concept that underscores the importance of unifying multiple ideas to develop a coherent understanding. KI, as emphasized by Linn and Eylon (2011), plays a crucial role in science learning and instruction. They propose leveraging the rich repository of ideas students already possess about any science subject and encouraging active engagement and inquiry. The KI approach also posits students benefit more from interacting with inquiry-based models and tools than from conventional lecture and textbook-based methods (Bransford et al., 1999; Linn & Eylon, 2006; Linn, Lee, Tinker, Husic, & Chiu, 2006). It values student-initiated inquiry and builds upon the ideas students develop independently. This approach argues that giving students control over their learning allows for more informative feedback about their understanding, particularly when they systematically investigate each variable and their interactions. For instance, a common misconception that metals are naturally cold can be addressed through the KI process. Eliciting students' pre-existing ideas is a crucial first step toward more nuanced understanding. By gathering these responses, the KI process benefits from a wide array of ideas, utilizing this collective knowledge to help students critically evaluate and refine their observations. The KI approach uniquely emphasizes an instructional design that recognizes the richness and diversity of students' ideas, prompting learners to assess new concepts, use evidence to compare differing viewpoints, and choose the most viable alternatives, all while fostering a continuous reflective process throughout their lifetime.

Linn and Eylon (2011) outline a general KI instructional pattern involving four processes: eliciting ideas, adding ideas, discovering, distinguishing ideas, and reflecting on ideas. This pattern informs the instructional design in this study, and is grounded in research on how students develop and refine their understanding of the natural world. It recognizes the richness and diversity of students' ideas and encourages them

Obaid et al. (2023b) also highlighted how learners frequently struggle to sustain conceptual links between disparate ideas, resulting in a fragmented understanding where learners fail to perceive the interconnectedness of various concepts. Furthermore, they noted a tendency among learners to retain invalid ideas even after exposure to correct ones, impeding the KI process by preventing learners from fully embracing new, accurate knowledge. Lastly, learners often struggle to build upon partial or incomplete links between ideas, limiting their ability to integrate knowledge.

Given the challenges identified by Obaid et al. (2003b), this study introduces a rule-formation tool within the context of simulation-based discovery learning. This tool is designed to function as a facilitator, aiding students in making sense of their observations, establishing connections with new ideas, distinguishing these from their existing concepts, and building upon partial links.

In line with this, the KI framework was utilized in our study to enhance the understanding of electric circuits. This framework guides students to reflect on their existing ideas and experiences, apply these ideas, and progressively integrate new knowledge. The aim is to leverage the rich repository of student’s ideas and promote their active engagement in the learning process. Within this theoretical context, the rule formation tool is envisioned as a facilitator, guiding students in making sense of their observations, connecting with new ideas, and distinguishing these from their existing concepts.

2.10. Bridging Gaps in the Literature

In this section, I highlight the gaps in the existing literature and establish the connections between these gaps and our research questions. To begin, the literature on discovery learning and simulation-based inquiry-based learning has provided valuable insights into learners’ ability to identify patterns and causal relationships independently, without strong guidance, and the transition from explicit instruction to exploration. However, there remains a debate regarding the efficacy and limitations of discovery learning, with differing views on its structure and guidance. These gaps in understanding effects of discovery learning, especially in simulation-based environments, form a foundational gap that this research seeks to address by examining the role of supports in rule discovery.

Moving on to the integration of engineering approaches in learning, the Next Generation Science Standards (NGSS) provide a framework for fostering critical thinking and innovation in students. While NGSS outlines engineering practices, there is a need to explore how these practices are integrated into simulation-based learning. This connection between NGSS and the integration of engineering approaches in simulation-based learning represents another significant gap that this study aims to fill by investigating how students navigate the engineering design process.

Search strategies in discovery learning have been extensively studied, uncovering various hypothesis testing strategies and their implications for learning. However, there is a need to delve deeper into how these strategies interact with the use of simulation and rule formation tools in a learning context. This gap underscores the relevance of this research, which explores students' engineering and search strategies in relation to hypothesis testing patterns within simulation-based learning.

Heuristics, often referred to as "rules of thumb," have shown promise in aiding complex problem-solving in discovery learning. However, the literature lacks a comprehensive exploration of the integration of heuristics within simulation-based discovery learning environments. This gap in understanding the effectiveness of heuristics in simulations aligns with this study's focus on heuristic principles and interventions in a simulation-based context.

Furthermore, the literature on selecting evidence to test hypotheses, a fundamental aspect of scientific discovery learning, highlights the importance of considering learners' positive test strategies and their exploration of alternative hypotheses. These insights reveal a gap in understanding how these strategies interact with engineering approaches and search strategies within simulation-based learning, a gap that this research seeks to bridge by examining distinct learning paths.

Rule formation when learners interact with computer models and simulations has been explored in various contexts, but there is limited research on its application in simulation-based discovery learning, particularly in the context of electric circuits. This gap in knowledge integration between rule formation and simulation-based learning informs my research focus on rule formation in electric circuits and its impact on students' understanding.

Self-regulated learning (SRL) plays a critical role in learner agency, yet there is a need to understand how SRL may be facilitated within simulation-based discovery learning. This gap in comprehending the relationship between SRL and rule formation in simulations aligns with this research, which emphasizes the active role of students in their learning process and the use of a rule formation tool to enhance SRL.

Lastly, the literature on knowledge integration (KI) underscores its significance in science learning but often lacks a detailed exploration of how KI can be supported within simulation-based learning environments. This gap in the application of KI principles within simulations aligns with my research's use of a rule formation tool to assist students in connecting and differentiating between existing and new ideas in the context of electric circuits.

In summary, the gaps identified in the literature highlight the need for a comprehensive investigation into how engineering approaches, search strategies, heuristics, hypothesis testing, rule formation, self-regulated learning, and knowledge integration interact within simulation-based discovery learning. My research questions address these gaps.

RQ 1. What progress in engineering approaches, search strategies, and integrating ideas do students make while using the electric circuit simulation in the control condition versus rule formation conditions?

RQ 2. What distinct learning paths do students take as they engage with the electric circuit simulation in the control condition versus rule formation conditions?

Chapter 3. Methodology

3.1. Participants

Data for this study were collected during the COVID-19 pandemic. I initially recruited a total of 71 undergraduate students from Simon Fraser University (SFU), Kwantlen Polytechnic (KPT), and the University of British Columbia (UBC). The study was advertised using university-wide email lists, with participants volunteering their time in return for $30 compensation. Ethics approval was obtained prior to recruitment from each of the aforementioned institutions. The selection process yielded a convenience sample, predicated on the participants' availability and willingness to engage, as opposed to a random sampling methodology. The pretest utilized a knowledge integration rubric, which classifies responses into six levels: 0 (blank response), 1 (off-task response), 2 (non-scientific idea), 3 (partial link), 4 (one valid link), and 5 (two or more valid links). In this study, participants scoring 0, 1, or 2 on the pretest were specifically included in the research sample. These scores indicate either no understanding or minimal understanding of series electric circuits and Ohm's law, suggesting a substantial gap in their foundational knowledge likely to have been introduced during high school education. Out of the 71 undergraduates initially recruited for the study, 60 participants met these criteria, demonstrating the significant educational need within this group. This observation emphasizes the notion that content widely expected to have been taught and learned in high school does not necessarily persist into undergraduate studies. Consequently, the selection process was intentionally designed to include these individuals, aiming to identify those who could most benefit from the educational interventions and thereby establishing a clear baseline for evaluating the interventions' impact on conceptual understanding. By strategically focusing on individuals demonstrating the greatest potential for conceptual growth, the study excluded participants with higher pre-test scores from further analysis. This approach led to a final sample of 60 students (N = 60) with low pretest scores being included in the data analysis, consisting of 40 students from UBC, 10 from SFU, and 10 from KPU. Participants were randomly assigned to three conditions of Control (n = 20), Experimental 1 (n = 20), and Experimental 2 (n = 20). The distinction between the two experimental groups lies in the intervention tools provided. Specifically, Experimental group 1 was given the Decision Table and Rule Induction intervention tool, while

Experimental group 2 solely received the Rule Induction intervention tool. Following the session, all participants were requested not to disclose experiment details to others.

3.2. Ethics

The ethics application, along with its accompanying document(s), underwent a thorough review by the Simon Fraser University Research Ethics Board (REB). The procedures were subsequently approved based on the ethical standards pertaining to research involving human participants. Each participant was provided with a comprehensive consent form, clearly outlining the purpose of the study, and they subsequently signed it to indicate their informed agreement to participate. The consent form informed participants about the use of a digital platform for the unit of instruction, the Web-based Inquiry Science Environment (WISE), which is hosted on a server at the University of California (UC) Berkeley and risks associated because data were stored outside Canadian borders subject to United States information protection laws.

3.3. Materials and Instrumentation

Participant behavior was recorded through a combination of a desktop screen recorder and a Zoom camera. The former captured voice recordings and on-screen interactions with the simulation, including note-taking activities, while the latter recorded facial expressions. Open broadcaster software was the chosen screen recorder software, and the researcher operated it.

3.3.1. Simulation-Activity Tasks

3.3.1.1. Terminology

Participants received concise definitions of key terms such as circuit, series vs. parallel circuits, current, resistor, resistance, voltage, voltage drop, voltmeter, and ammeter. They could revisit these terms at any point during the session by clicking on the WISE screen's "Refresh your memory" button.

3.3.1.2. Pre and Post Tests

Six pretest and six identical posttest questions were administered, with a total of 10 minutes allowed for answering all six questions (see Appendix A). Questions about current and voltage drop in series electric circuits were presented to participants via the WISE platform. The format of each question followed a similar structure:

1. Presentation of a scenario or a technical description: Scenario questions began with a scenario describing the creation and testing of circuits. For instance, one scenario described a person named Cannice who created Circuit A and observed a particular outcome related to voltage drop. Another person named Amy posed a question or inquiry about a different, related circuit (Circuit B). The initial character then draws a conclusion based on their earlier observations or evidence. Technical questions provided a technical description of a circuit setup, detailing component attributes and posing a question about the expected behavior of the circuit.
2. Display of circuit diagrams: Circuit diagrams were displayed under each question.
3. Multiple choice question (MCQ) format: Participants were instructed to select the best answer from three options, labeled A, B, and C. For instance, in the scenario involving Cannice and Amy, the options were: A. Cannice's conclusion is correct; B. Cannice's conclusion is incomplete; and C. Cannice's conclusion is incorrect.
4. Short answer explanation: Immediately following the MCQ, participants were asked to explain their selected choice of A, B, or C.

3.3.1.3. Orientation Task

Participants were instructed to respond to questions on the WISE platform and were given an instructional video with an overview of the electric circuit simulation interface before they interacted with the simulation. This was intended to reduce the cognitive load of interacting with the novel environment.

3.3.1.4. Electric Circuit Tasks

The tasks were structured in three phases: exploration, prediction, and investigation. The nature of the tasks varied based on these phases, and detailed instructions were provided to guide participants through each phase, including exploring basic relationships, predicting outcomes, investigating simulations, and using tools.

1. Exploration Phase: All participants were asked to investigate a schematic of a

simple series circuit including one battery and one bulb. Using a simulation, they were tasked with replicating the circuit and determining the relationships between resistance and current, voltage and current, and resistance and voltage drop. Before progressing to the subsequent phase, the tutor verified that each participant had accurately identified and understood these relationships.

2. Prediction Phase (without simulation):
    a. Control Group: Participants were instructed about the goal (Figure 3.1) and how to use the provided drawing tool shown (Figure 3.2). They were guided to create circuits and predict the corresponding changes in current and voltage drop. The drawn circuits were to be added by the participants to the WISE environment notebook (Figure 3.3). Their subsequent task involved writing as many predictions as possible about the current and voltage drop for each drawn circuits, following a specific format as shown in Figure 3.4.

Figure 3.1 Control Group Prediction Phase: Goal Instructions

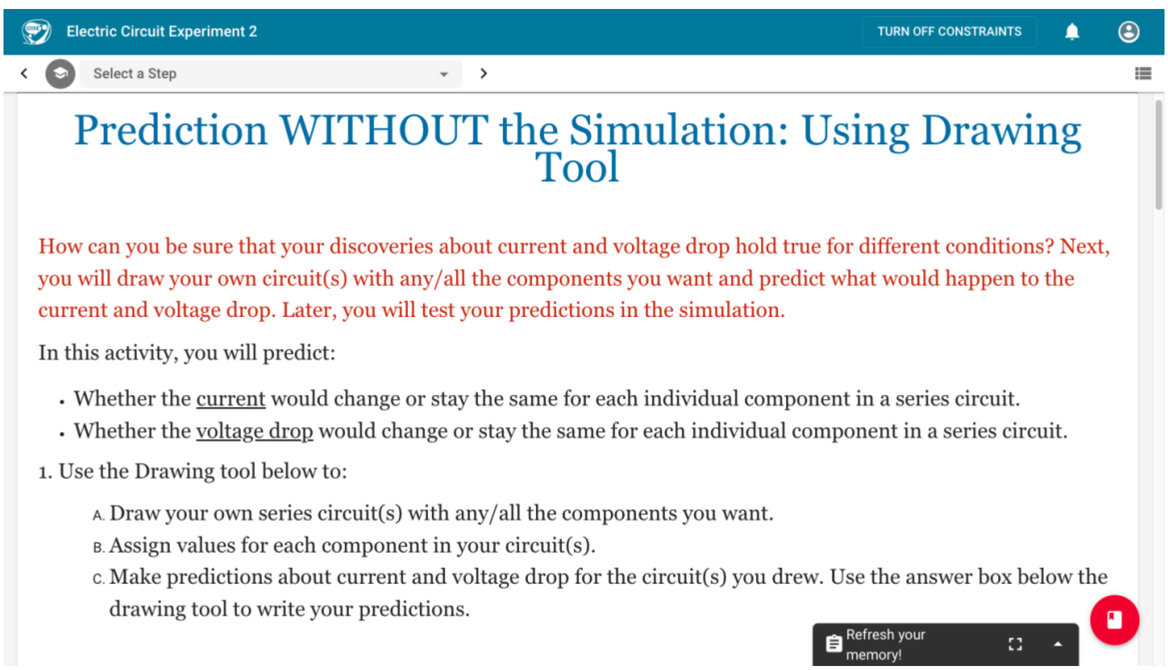


Figure 3.2 Control Group in Prediction Phase: Drawing Tool Instruction

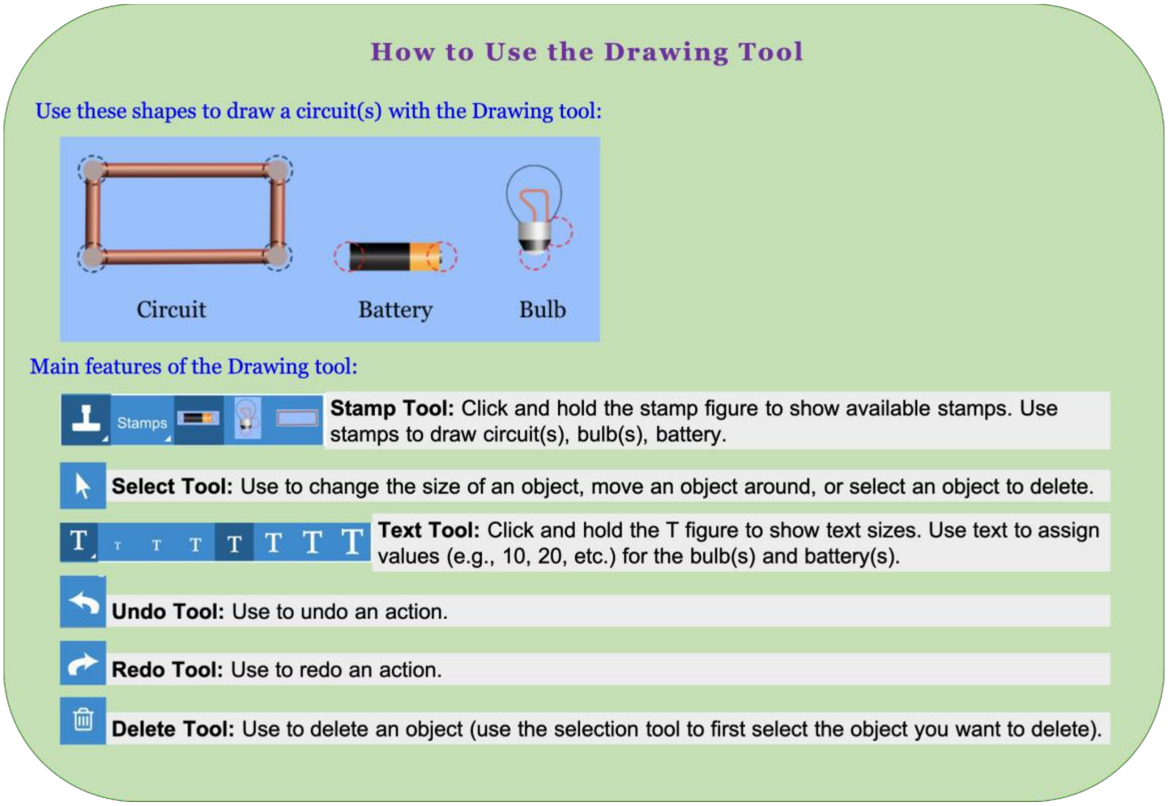


Figure 3.3 Control Group in Prediction Phase: Add to Notebook

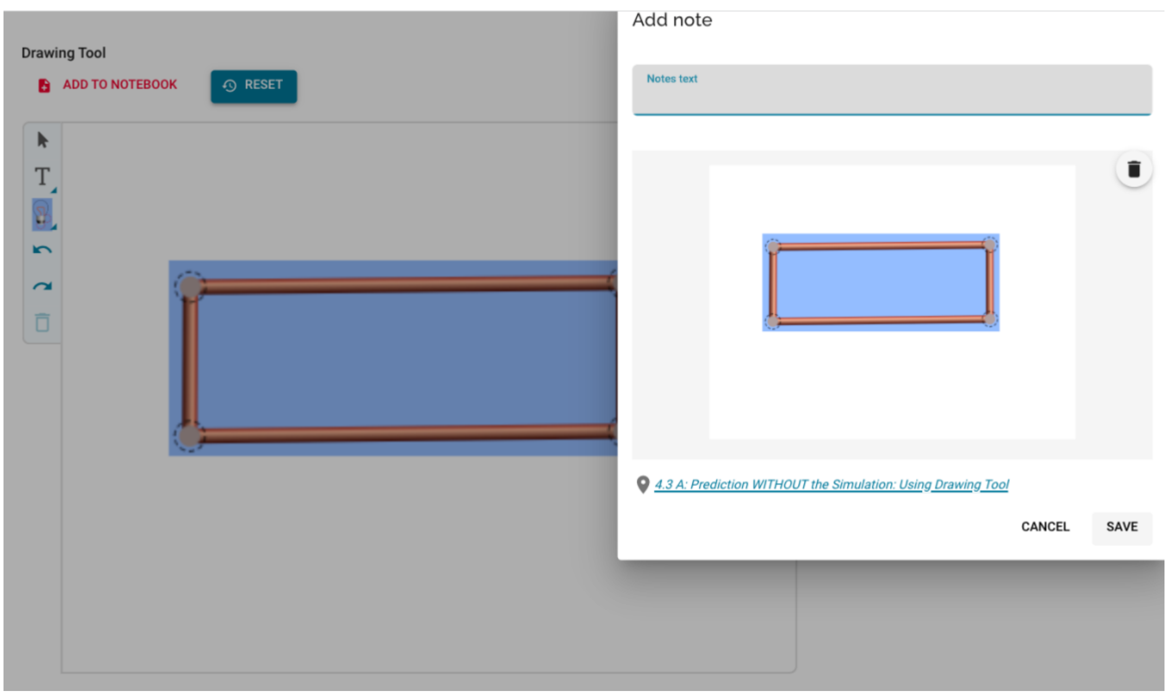
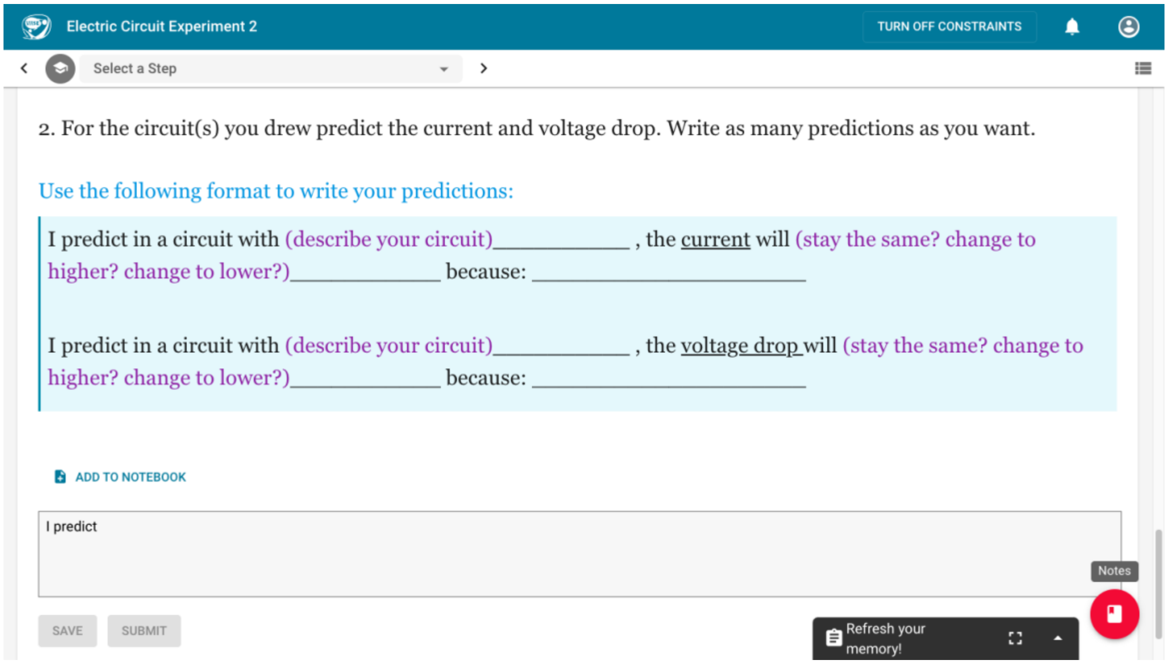


Figure 3.4 Control Group in Prediction Phase: Subsequent Task



b. Decision Table and Rule Induction Intervention Group (Experimental group 1): This group was briefed on the goal (Figure 3.5), and examples were provided to devise comparative trials (Figure 3.6) and generate rules (Figure 3.7). They were given an example of an IF-THEN rule with highlighted variables for current and voltage drop (Figure 3.8). Their task involved using the given empty tables to create as many comparative trials as possible by identifying constant and changing variables (Figure 3.9). They added these comparative trials to the WISE environment notebook, prepared by the researcher.

Figure 3.5 Experimental Group 1 in Prediction Phase: Goal Instruction

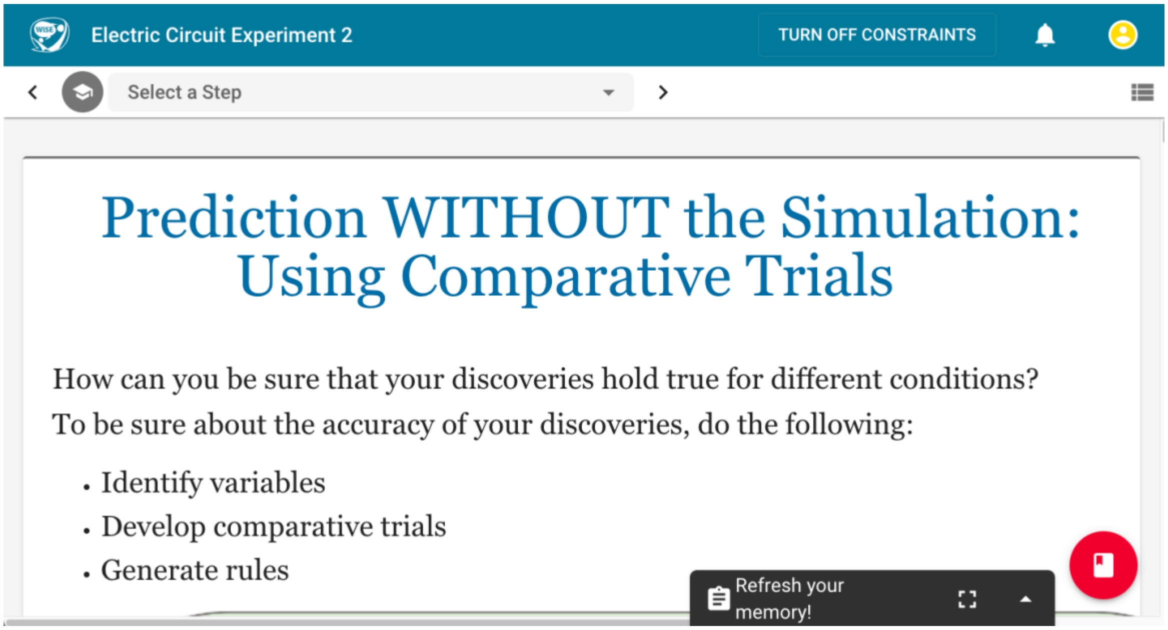


Figure 3.6 Experimental Group 1 in Prediction Phase: Example of Comparative Trials

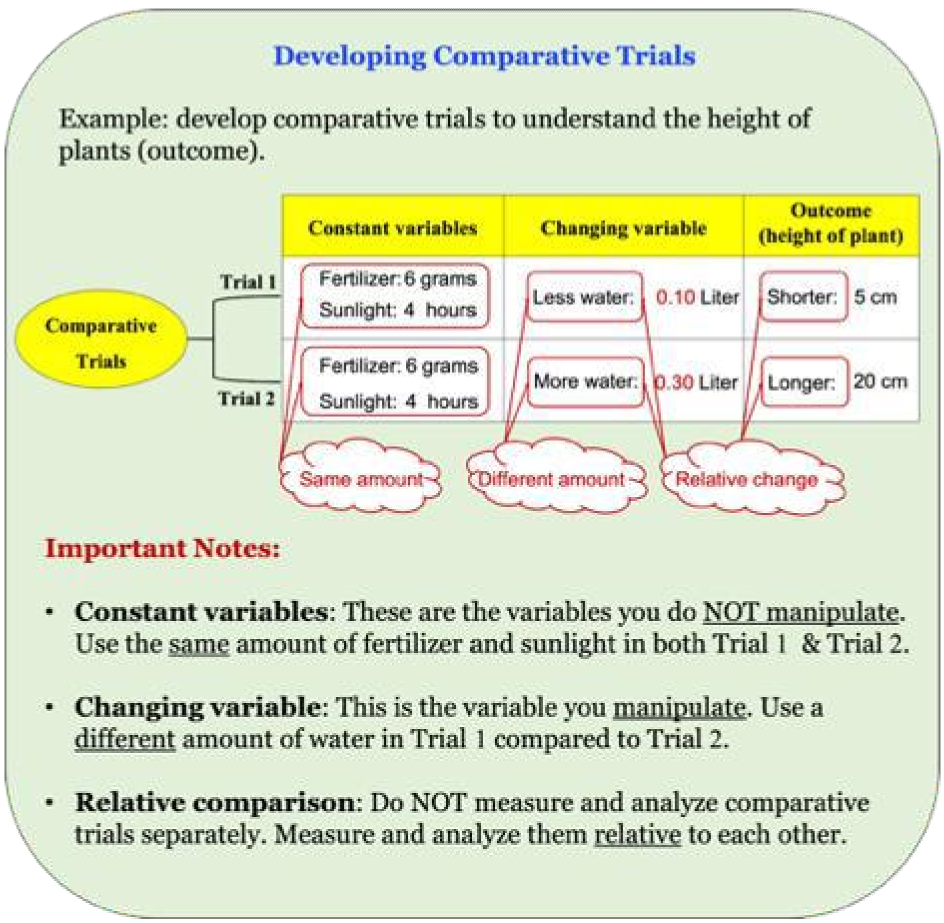


Figure 3.7 Experimental Groups in Prediction Phase: Example of Generating Rules

Generating Rules

Rules determine how different variables interact with each other. Develop rules for your comparative trials.

A rule consists of 3 parts:

(1) FOR ALL CONSTANT VARIABLES (variables you don't change):

(e.g., 6 grams fertilizer & 4 hours sunshine in both Trial 1 & Trial 2)

(2) IF (variable(s) you change):

(e.g., 0.1 Liter water in Trial 1 compared to 0.30 Liter water in Trial 2)

(3) THEN (outcome):

(e.g., shorter plant in Trial 1 compared to longer plant in Trial 2)

This means: for these constant variables (6 gams fertilizer; 4 hours sunlight) IF we increase the amount of water from 0.1 Liter to 0.30 Liter, THEN we will get a longer plant.

Figure 3.8 Experimental Groups in Prediction Phase: Example of Variables

Current & Voltage Drop Discoveries

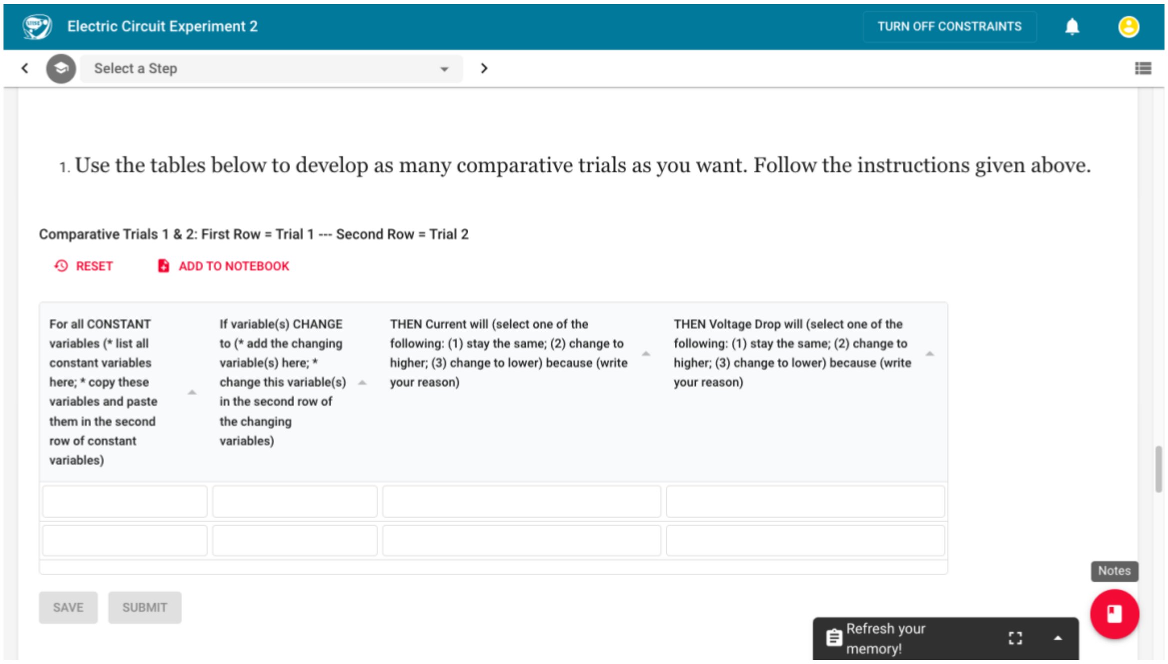
The colored boxes below represent different variables. You may use the main components (battery, bulb, ammeter, and voltmeter) differently to understand what happens to the current and voltage drop.

IF we have a series circuit with 1 battery with 10 voltage and 1 bulb with 10 resistance and an ammeter placed after the bulb and a voltmeter placed across the bulb, THEN the Current (outcome) will be 1.00 Ampere and the voltage drop (outcome) will be 10 volts.

Your goal is to use different variables to predict these outcomes:

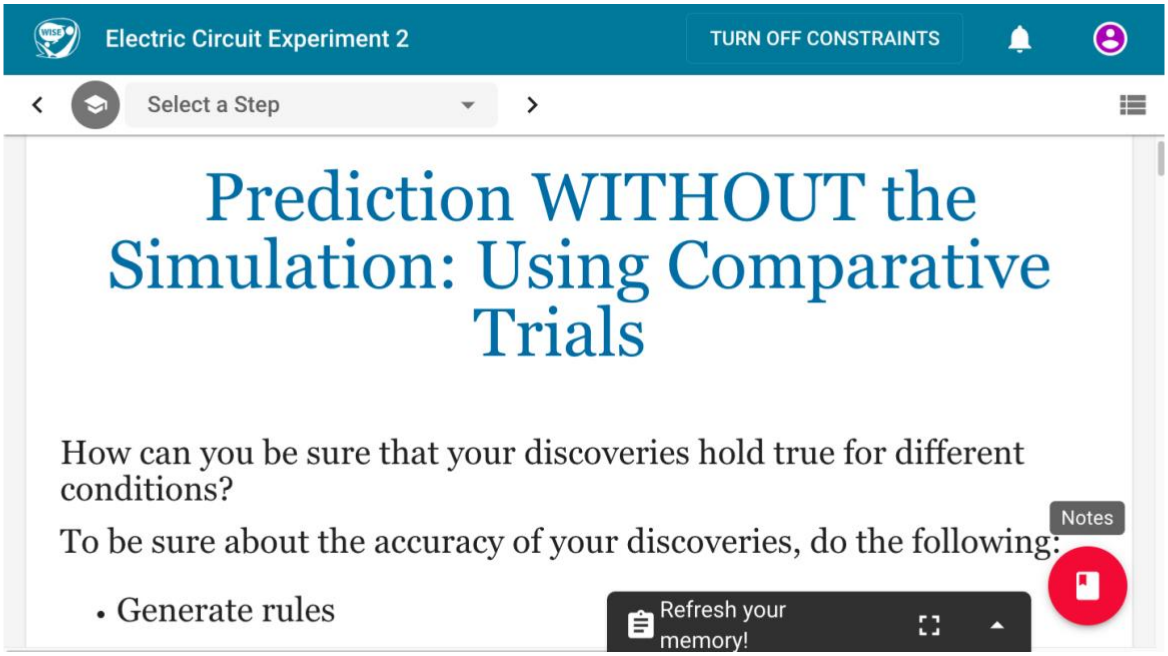
- Whether the current would change or stay the same for each individual component in a series circuit.
- Whether the voltage drop would change or stay the same for each individual component in a series circuit.

Figure 3.9 Experimental Groups in Prediction Phase: Subsequent Task



c. Rule Induction Intervention Group (Experimental group 2): This group was briefed on the goal (Figure 3.10) and examples were provided to generate rules (Figure 3.7). Mirroring the other intervention group, they were offered an example of an IF-THEN rule regarding current and voltage drop, with the variables highlighted for clarity (Figure 3.8). Upon reviewing these examples, they were instructed to develop numerous comparative trials by discerning the constant and changing variables in order to predict current and voltage drop (Figure 3.9). They added these comparative trials to the notebook within the WISE environment.

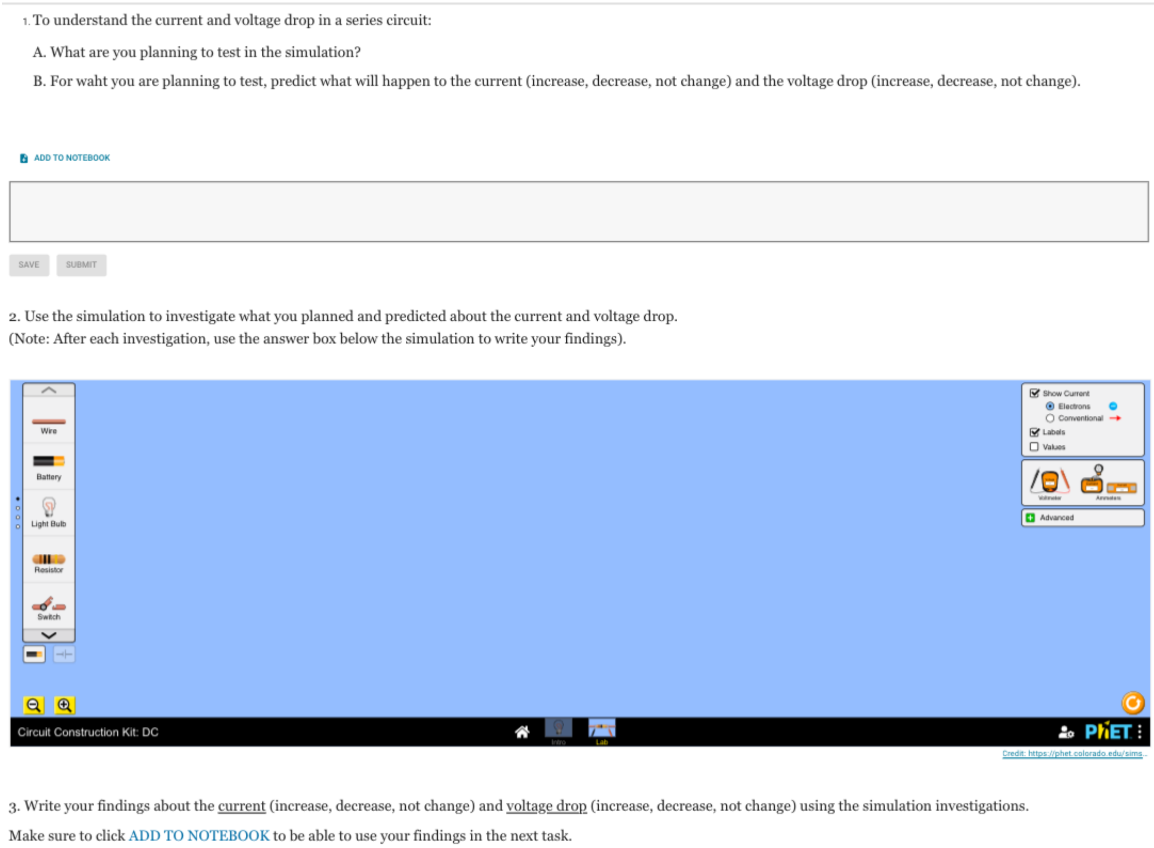
Figure 3.10 Experimental Group 2 in Prediction Phase: Goal Instruction



3. Investigation Phase:

    a. Control Group: Participants were prompted to outline their simulation testing plans and predict the possible changes in current (increase, decrease, no change) and voltage drop (increase, decrease, no change) (Figure 3.11). They then used the simulation to test these predictions, with the findings recorded in the WISE environment notebook.

Figure 3.11 Control Group in Investigation Phase: Tasks



1. To understand the current and voltage drop in a series circuit:

A. What are you planning to test in the simulation?

B. For waht you are planning to test, predict what will happen to the current (increase, decrease, not change) and the voltage drop (increase, decrease, not change).

ADD TO NOTEBOOK

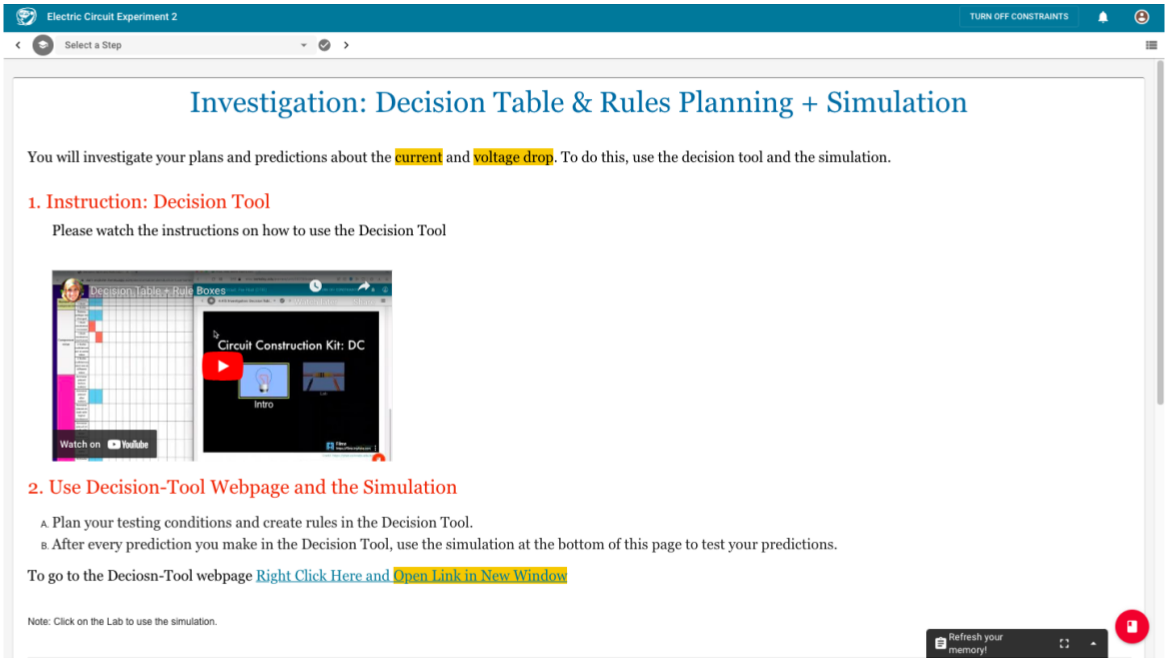
SAVE SUBMIT

2. Use the simulation to investigate what you planned and predicted about the current and voltage drop.
(Note: After each investigation, use the answer box below the simulation to write your findings).

3. Write your findings about the current (increase, decrease, not change) and voltage drop (increase, decrease, not change) using the simulation investigations.
Make sure to click ADD TO NOTEBOOK to be able to use your findings in the next task.

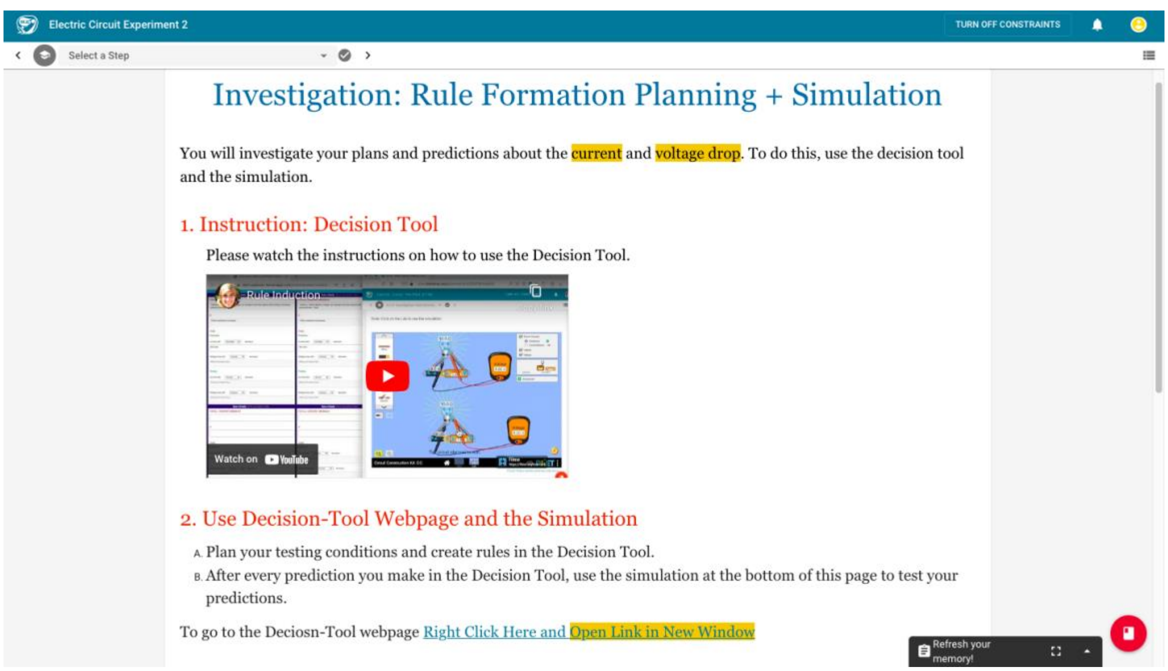
b. Decision Table and Rule Induction Intervention Group (Experimental group 1): Participants were instructed to examine their plans and predictions using the Decision Tool and the simulation (Figure 3.12). They were provided with an instructional video demonstrating how to utilize the Decision Table and Rule Induction Tool in tandem with the simulation. They were then asked to use the Decision-Tool Webpage and the simulation to plan their testing conditions, establish rules, and test their predictions.

Figure 3.12 Experimental Group 1 in Investigation Phase: Tasks



c. Rule Induction Intervention Group (Experimental group 2): Much like the previous group, these participants were also given an instructional video, guiding them on how to use the Rule Induction Tool and the simulation together (Figure 3.13). They were then asked to use these tools to plan their testing conditions, create rules, and test their predictions.

Figure 3.13 Experimental Group 2 in Investigation Phase: Tasks

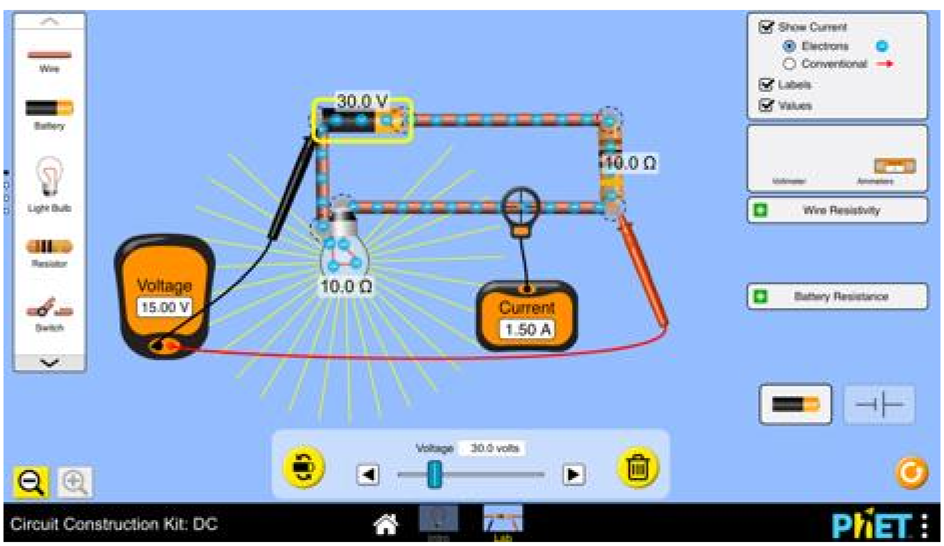


3.3.2. Simulation-Activity Tools

3.3.2.1. Simulation Tool

All participants were provided with a physics simulation sourced from the PhET free online collection of educational simulations, which was embedded in the WISE website (Figure 3.14). This simulation (https://phet.colorado.edu/en/simulation/circuit-construction-kit-dc), enabled students to investigate DC electric circuits using key components such as batteries, bulbs, resistors, wires, voltmeters, and ammeters. It included all the necessary items for circuit construction and a workspace that allowed participants to drag, drop, and interconnect the items to form various circuits. Moreover, the simulation facilitated taking measurements critical to understanding relationships involving current, resistance, and voltage drop.

Figure 3.14 A Sample of Electric Circuit Construction in PhET Simulation



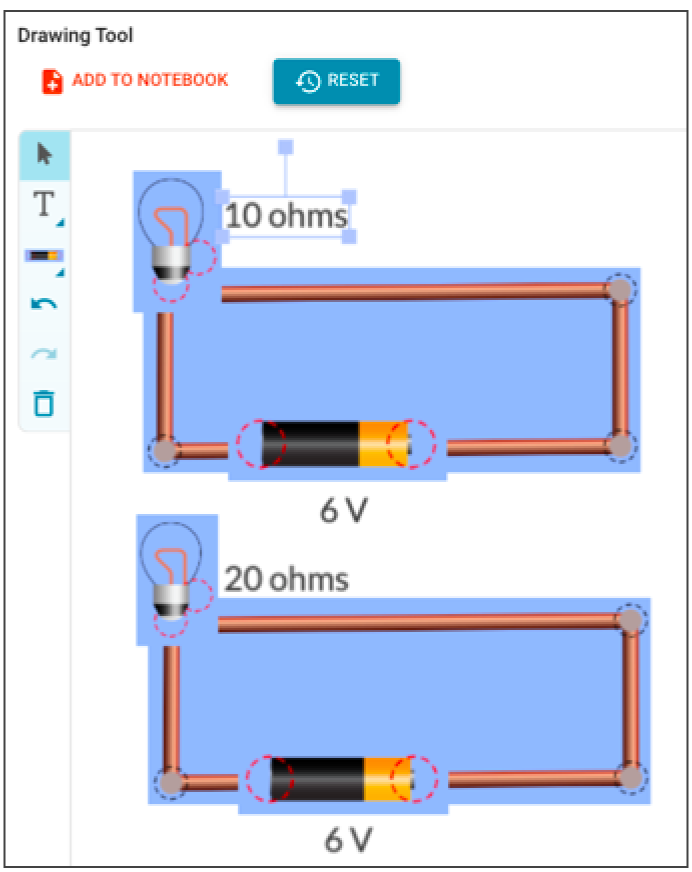
3.3.2.2. Drawing Tool

The Drawing Tool was utilized during the prediction phase for the Control group (Figure 3.15). Instead of relying on direct simulation results, students were prompted to draw a series of electrical circuits using any or all components (such as wires, batteries, and bulbs) and predict the ensuing current and voltage drop. They were then instructed to add these drafted circuits to a notebook (Figure 3.3) in the WISE environment.

1. Stamp Tool: Students could choose from available circuit components such as wires, batteries, and bulbs and place these onto the workspace for the construction of personalized circuits.
2. Select Tool: Students could select specific components after positioning them in the circuit workspace, relocate them, and resize them by clicking and dragging one end of the component.
3. Text Tool: By selecting the text size and clicking on the workspace, students could assign values for each component.
4. Undo/Redo Tools: These tools provide corrective options during the design process. The Undo Tool allowed students to reverse previous actions, and the Redo Tool reinstated undone actions.
5. Delete Tool: This tool allowed removing unwanted components from the design.

For each devised circuit, students were required to note predictions in a designated answer box (Figure 3.4).

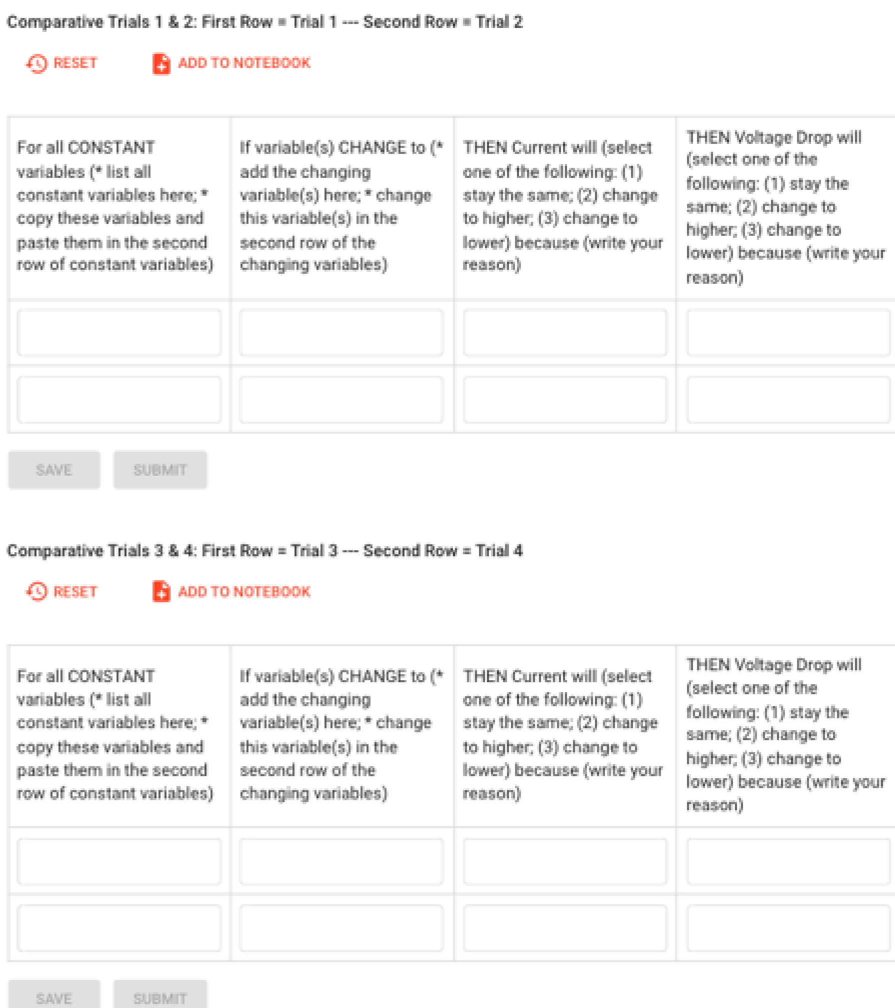
Figure 3.15 Control Group in Prediction Phase: Drawing Tool



3.3.2.3. Simple Tables

This tool was utilized during the prediction phase for both intervention groups (Figure 3.16). Without returning to the simulation, students were prompted to generate as many comparative trials as possible using provided tables. Each table encompassed four columns, designated for constant variables, changing variable(s), current prediction and reasons, and voltage drop prediction along with reasons. Each comparative trial consisted of two rows, corresponding to a trial's compared design respectively.

Figure 3.16 Experimental Groups in Prediction Phase: Comparative Tables



3.3.2.4. Decision-Based Automated Rule-formation Tool (DART)

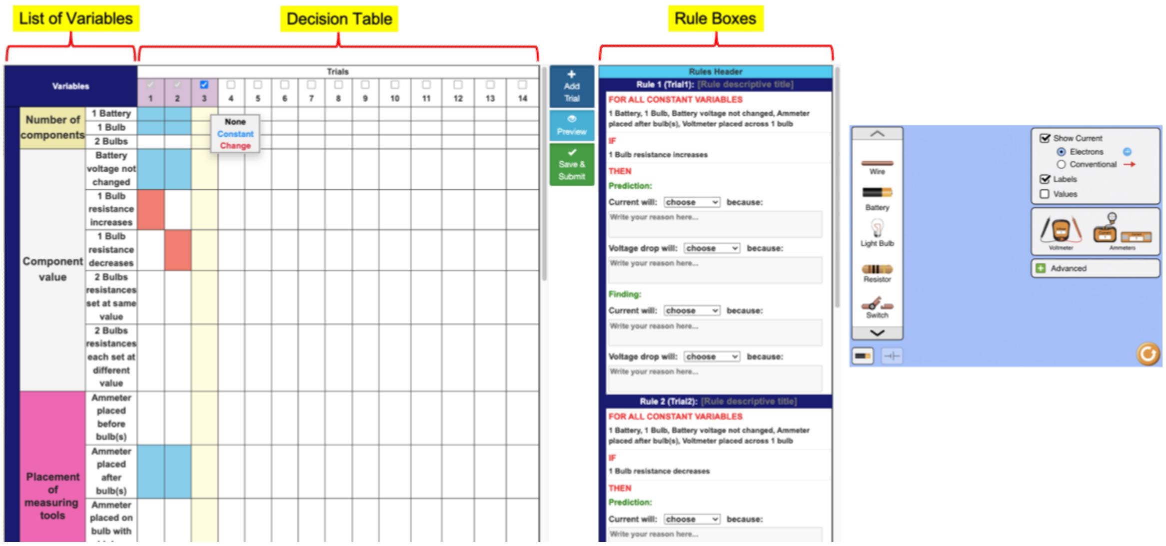
I conceptualized and developed this tool to offer heuristic support in the process of discovery and to generate data regarding students' selective search methodologies (Figure 3.17). DART offers several distinct features:

1. A List of Relevant Variables: This feature lists all variables that could be manipulated in designing a circuit.
2. A Decision Table: In this table, students craft combinations of variables to design trials and identify which variables in comparative trials remain constant and which ones differ.
3. Rule Boxes: These boxes automatically generate a IF-THEN rule as students develop a decision table. They identify parameters of a rule: 'For all constant variables', 'IF', 'Then (Prediction and Findings)'. The prediction and finding boxes within the rule box contain drop-down menus from which students can select values to describe the behaviors of current and voltage drop. Additionally, for the prediction and finding boxes, students are provided a space to document their reasoning.

As students progress through various trials employing the rule-formation tool, they engage in an iterative hypothesis testing process. Each comparative trial offers students a chance to test their hypotheses by choosing constant and changing variables, then making predictions regarding current and voltage drop. By comparing their predictions with the actual findings, students gather evidence to evaluate the validity of their hypotheses.

The heuristic support supplied by this tool could underscore students' uncertainties and enable them to engage in rule-based selective searching, a process often performed ineffectively by most adults (Wason, 1960, 1983).

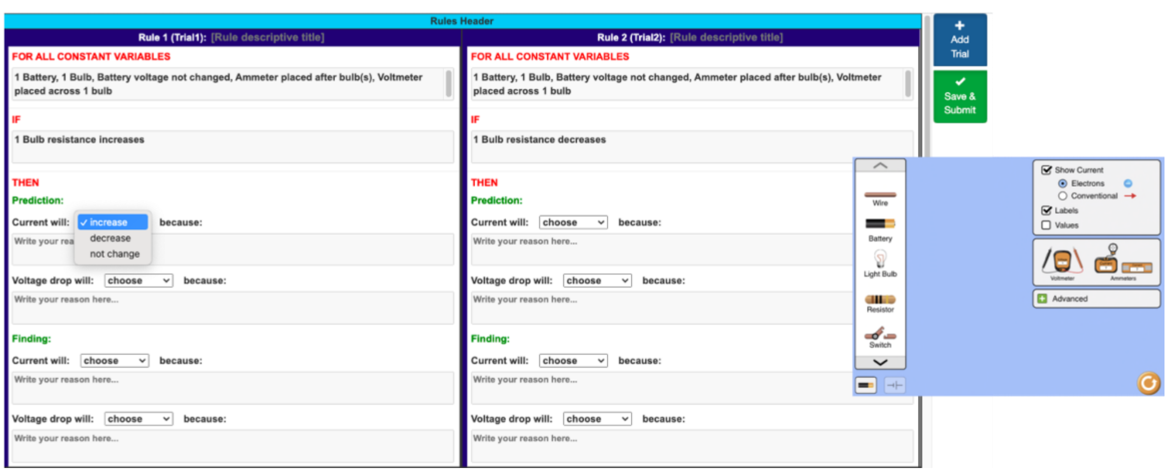
Figure 3.17 Experimental Group 1 in Investigation Phase: Decision Table + Induction Rule



3.3.2.5. Rule-Formation Tool (RT)

During the investigation phase, students involved in Experimental group 2 employed a rule formation tool to establish IF-THEN rules (Figure 3.18). This group, unlike experimental group 1, was required to identify the variables for the comparative trials independently. A part of the tool labeled 'FOR ALL CONSTANT VARIABLES' denoted the control conditions. Other sections labeled 'IF' and 'THEN' provided students with the capacity to detail a hypothesis in terms of a prediction and findings. The prediction and finding fields in the tool included drop-down menus from which students could select values pertaining to current and voltage drop.

Figure 3.18 Experimental Group 2 in Investigation Phase: Induction Rule



3.3.3. Coding

The coding process involved analyzing students' responses in the pre- and post-test according to a knowledge integration rubric. The rubric assigns scores from 0 (blank response) to 5 (multiple valid links): off task = 1; non-normative/irrelevant terminology use = 2; partial = 3; 1 valid link = 4; 2 or more valid links = 5.

I coded participants’ engineering approaches to selecting variables and search strategies by inspecting their designed comparative trials simulated during the investigation phase. Each comparative trial compared two designs in which values for variables were selected relative to each other.

To represent engineering approaches (i.e., variable selection), I developed a dictionary. Each entry within the dictionary represented a different engineering approach, and each approach contained a nested dictionary of comparative trials. Within each comparative trial, attributes reflected the specific variables chosen for that trial. The dictionary's structure is outlined in Figure 3.19. An example of how instances, comparative trials, and attributes might be populated is as follows in Figure 3.20.

Figure 3.19 Engineering Approaches Hierarchical Structure

```
engineering_approaches (dictionary)
    |
    --> Instances (keys)
        |
        --> Comparative trials (nested dictionaries)
            |
            --> Attributes (values for keys)
```

Figure 3.20 Coding Engineering Approaches

```
engineering_approaches = {
    "Instance 1: Bulb": {
      "Comparative trial: equal": [
        "Attribute 1: 1 bulb",
        "Attribute 2: 2 bulbs"
      ],
      "Comparative trial: different": ["Attribute 1: 1 bulb vs. 2 bulbs"],
    },
     "Instance 2: Battery Voltage": {
      "Comparative trial: equal": [],
      "Comparative trial: different": [],
    },
    "Instance 3: Resistance": {
      "Comparative trial: equal": [
        "Attribute 1: 1 bulb",
        "Attribute 2: 2 bulbs different value",
        "Attribute 3: 2 bulbs (or 1 vs. 2) same value",
        "Attribute 4: 2 bulbs different value with changed position",
      ],
      "Comparative trial: different": [
        "Attribute 1: 1 bulb or 1 vs. 2 bulbs",
        "Attribute 2: 2 bulbs"
      ],
    },
    "Instance 4: Ammeter": {
      "Comparative trial: equal": [
        "Attribute 1: on/after/before 1 or 2 bulb",
        "Attribute 2: on high/low of 2 bulbs"
      ],
      "Comparative trial: different": [
        "Attribute 1: before vs. after 1 or 2 bulbs",
        "Attribute 2: on high vs. low of 2 bulbs"
      ],
    },
      Instance 5: Voltmeter": {
      "Comparative trial: equal": [
        "Attribute 1: across 1 or 2 bulbs or battery",
        "Attribute 2: across high/low resistance of 2 bulbs"
      ],
      "Comparative trial: different": [
        "Attribute 1: across high vs. low resistance of 2 bulbs (or 1 vs. 2)",
        "Attribute 2: across high vs. low resistance of 2 bulbs vs. battery",
    }
}
```

In analyzing the participants' data, a set of codes was developed by the researcher to categorize various search strategies. These codes, presented in Table 3.1, were used to systematically categorize each search strategy evident in the data. While the majority of the codes were derived based on the observed patterns and nuances of the participants' actions, the Rule Generation category specifically drew inspiration from Bruner's discovery search strategies.

Table 3.1 Codes for Search Strategies

Category	Definition	Codes
Action Variables	Taking actions such as making predictions, performing confidence measures, identifying a problem, and generating rules for current and voltage drop in each comparative trial. Initiating distinct comparative trials in the investigation phase compared to the exploration phase	● No action ● New comparative trial ● No new comparative trial
Prediction	Predicting the current & voltage drop for each comparative trial	● Use same rule from preceding rule ● Fill up gaps from preceding rule
Confidence Measure	Verifying/falsifying prediction of current & voltage drop after testing comparative trials in simulation	● Falsify prediction ● Verify prediction
Problem Identification	Identifying unexpected outcomes in current & voltage drop and finding ways to change the outcome	● Identify a problem
Rule Generation	Encountering the sequences of confirming and/or infirming contingencies in the previous rules to form new rules for the current & voltage drop	● Confirmation redundancy ● Successive scanning ● Simultaneous scanning ● Conservative Focusing ● Focus gambling

Two raters, both experienced in the application of the KI rubric across multiple science-related research studies, independently coded the data. Prior to coding, they engaged in a series of discussions to ensure a comprehensive understanding of the rubric and the coding scheme for engineering approaches and search strategies. During these discussions, they also clarified any ambiguities in the rubric and made any necessary revisions to ensure consistency in coding. After independent coding, the two raters compared their results to identify discrepancies and discussed any disagreements until a consensus was reached. To quantify the consistency between the raters, inter-rater reliability was calculated using Cohen's kappa coefficient. A strong inter-rater reliability was achieved, $\kappa = 0.81$.

3.3.3. Data Analysis

3.3.3.1. Levenshtein Edit Distance

I utilized Levenshtein edit distance to analyze sequences of student engineering approaches, search strategies, and ideas generated in post-test. The Levenshtein edit distance is a metric that measures the degree of similarity between two strings of text or sequences of ideas, making it an ideal method for this investigation. You might think of the Levenshtein edit distance as a measure of how much effort it would take to transform one string of ideas into another. In applying this method, I adopted a unit cost for operations, meaning each operation—whether a replacement, insertion, or deletion—had an equal cost of 1. This choice was made to simplify the analysis while capturing the essence of the students' strategy shifts. This method has been widely applied in fields such as linguistics and bioinformatics to identify similar strings or sequences.

The adoption of the Levenshtein edit distance as the metric for analyzing similarity was driven by the unique nature of the dataset—sequences of actions and responses—which are not adequately handled by traditional distance measures such as Euclidean or Manhattan distances. These conventional measures are well-suited for numerical or categorical data but fall short when dealing with the intricacies of sequence comparison (Géron, 2019). An alternative approach could have been the use of other sequence comparison algorithms like the Hamming distance. However, the Hamming distance is limited to sequences of the same length, making it less versatile than the Levenshtein edit distance for this study's varied sequence lengths (Navarro, 2001). The Levenshtein distance, capable of comparing sequences of different lengths and considering the order of elements, offers a more suitable and nuanced measure for this study's requirements.

Consider the following example in Table 3.2, which compares the search strategy strings of Student 1 and Student 2:

Table 3.2 Sample Sequences of Search Strategies for Student 1 and Student 2

Student 1	Student 2	Levenshtein edit distance
New comparative trial	New comparative trial	-
Prediction Current: Fill up gaps	Prediction Current: Fill up gaps	-
Prediction Voltage Drop: same rule	Prediction Voltage Drop: same rule	-
Confidence Current: verify prediction	Confidence Current: falsify prediction	Replace
Confidence Voltage Drop: verify prediction	Confidence Voltage Drop: falsify prediction	Replace
No action	No action	-
Rule Current: Successive scanning	Rule Current: Successive scanning	-
Rule Voltage Drop: Successive scanning	Rule Voltage Drop: Simultaneous scanning	Replace
Post-test: Voltage Drop partial	Post-test: Voltage Drop non-normative	Replace
Post-test: Current partial	Post-test: Current partial	-

From this comparison, we can calculate a Levenshtein edit distance of 4 between the strings of Student 1 and Student 2. This distance represents the minimum number of replacements required to transform the search strategy string of Student 1 into that of Student 2. Such analysis enables us to quantitatively compare and contrast the sequence of search strategies employed by different students in the study.

3.3.3.2. Sequence Clustering using K-Means and Levenshtein Edit Distance

In this study, sequence clustering was performed using the K-means algorithm provided by the scikit-learn library (version 0.23) in Python 3.8. I utilized K-means clustering of the Levenshtein edit distances to group participants according to similarities in the sequences of their engineering approaches, search strategies, and KI posttest responses. Clusters exhibit represent similarity in the values of these variables.

The K-means clustering algorithm was chosen for its proven efficacy in partitioning data into clusters, such that each data point belongs to the cluster with the nearest mean (Géron, 2019). Alternatives such as hierarchical clustering were considered. Hierarchical clustering, while advantageous for its dendrogram output providing a visual representation of cluster formations, does not necessitate specifying the number of clusters a priori (Géron, 2019). Compared to this methods, K-means demands a predetermined cluster count, offers a balance of simplicity, computational efficiency, and the flexibility to adapt to the use of non-traditional distance measures such as the Levenshtein edit distance.

K-means clustering begins by randomly assigning initial centroids to define clusters. These centroids are then iteratively adjusted by redistributing data points based on their distance to the current centroids and recalculating the centroids according to the updated cluster members. This process continues until a steady state is reached where centroids no longer shift.

In the conventional implementation of the K-means algorithm, similarity measures such as Euclidean or Manhattan distance are employed, which are particularly suited for numerical or categorical data. However, given the distinctive nature of the dataset in this study— sequences of actions and responses from participants—I adopted the Levenshtein edit distance as the metric to be analyzed for similarity. The application of Levenshtein edit distance in K-means clustering varies from the use of Euclidean or Manhattan distances. In the traditional numeric-based K-means clustering, a centroid is defined as a central location within each cluster and is initially assigned at random. Each data point is then ascribed to the cluster whose centroid is the nearest, using a distance measure such as Euclidean distance. However, when deploying K-means clustering on sequence data and utilizing Levenshtein distances, the concept of a centroid is adapted. Here, a "centroid" retains its role as a central point within each cluster, but is defined as a representative sequence that best encapsulates or summarizes the sequences within that cluster. The initial centroids were determined automatically by the K-means algorithm, without manual intervention, based on the Levenshtein edit distances among the sequences. During the assignment phase, each sequence in the dataset is compared to these centroid sequences through the Levenshtein distance. A sequence is assigned to the cluster whose centroid has the smallest Levenshtein distance to the sequence—that is, the centroid sequence that requires the fewest edits to transform into the dataset sequence. In the subsequent update phase, centroids are recalculated after all sequences have been ascribed to clusters. This entails identifying a new representative sequence for each cluster that minimizes the aggregate Levenshtein distance to all other sequences within that cluster. This new centroid could either be an existing sequence from the dataset or a novel sequence derived from cluster sequences, frequently determined through a process known as multiple sequence alignment.

This cycle of assigning sequences to clusters and recalculating centroids is iterated until the centroids cease to change, or the magnitude of change falls beneath a predefined threshold. Although this approach effectively employs the strengths of Levenshtein distance for sequence comparison in K-means clustering, it is noteworthy that it can be more computationally intensive due to the complexities involved in calculating Levenshtein distances and recalibrating the centroids.

3.3.3.3. Optimizing Cluster Numbers using Silhouette Coefficient

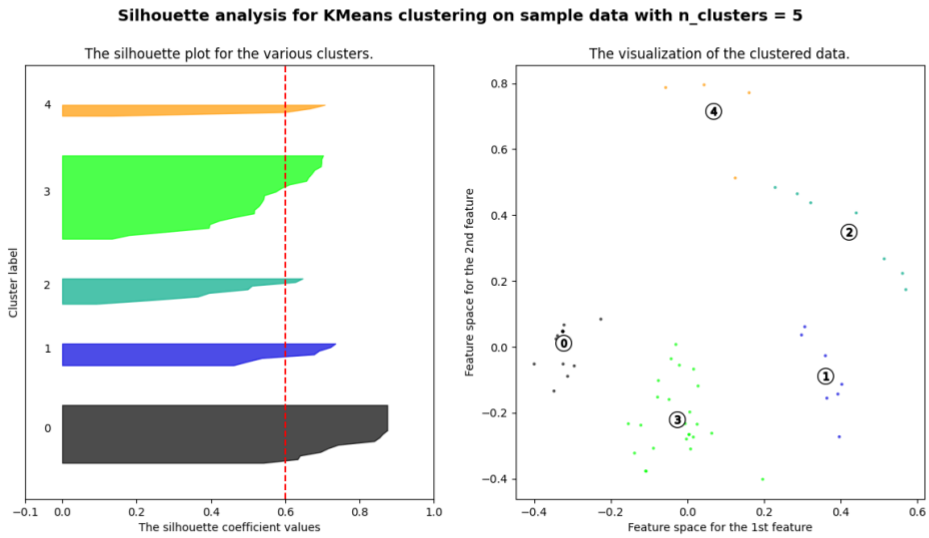
For optimal clustering, determining the ideal number of clusters is crucial. There are several techniques to determine this number, such as the Elbow Method, the Gap Statistic, and the Silhouette Coefficient. The Elbow Method and Gap Statistic are often used, but they can sometimes provide ambiguous results, depending on the structure of the data (Bies et al., 2006). In contrast, the Silhouette Coefficient provides a more consistent and robust measure of cluster quality, leading to its selection for our study (Géron, 2019).

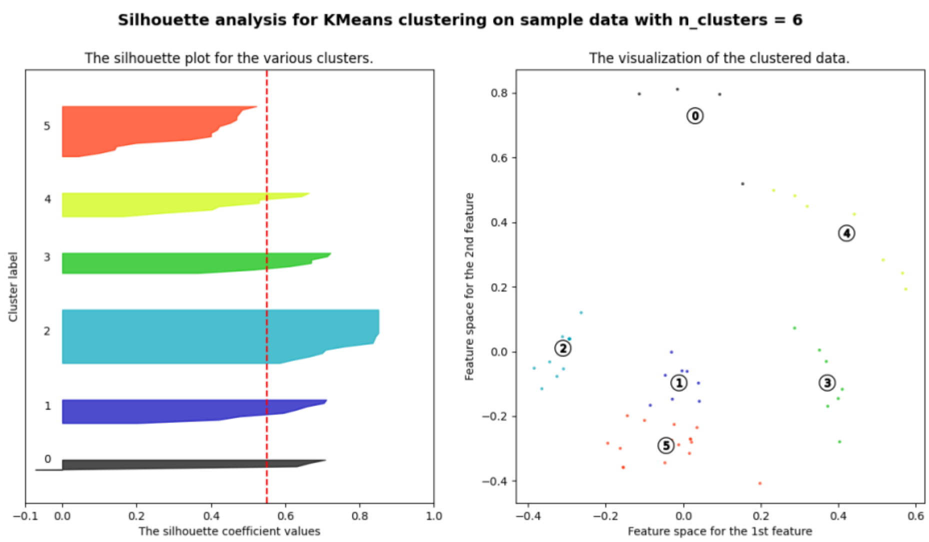
The Silhouette Coefficient computes the average distance between each data point and all other points within the same cluster (intra-cluster distance, denoted as ‘a’) and compares it with the average distance between each data point and all other points in the nearest different cluster (inter-cluster distance, denoted as ‘b’). Here, nearest different cluster means the cluster whose average distance to a data point is the smallest, but it's not the cluster to which the data point is currently assigned. The Silhouette Coefficient (s) for a single sample is then calculated as follows: s = (b - a) /

max(a, b). This value is calculated for each data point in the dataset and then averaged to determine the Silhouette Coefficient for the entire dataset. The Silhouette Coefficient ranges from -1 to +1. A high value near +1 indicates that the data point is well-matched to its own cluster and poorly matched to neighboring clusters, signifying a well-defined cluster structure and a suitable choice of the number of clusters.

In visualizing the Silhouette Coefficient a silhouette plot, provides a graphical representation of each cluster. Each data point is represented by a line, with its length determined by the Silhouette Coefficient value for that datum. This plot depicts the extent to which each data point matches its own cluster compared to other clusters. It helps to visually assess the quality of clustering, the cohesion within clusters, and the separation between them (Figure 3.21). As illustrated in Figure 3.21, in the silhouette diagram for n-clusters = 5 and n−clusters = 6, each cluster is represented by a silhouette shape. The height of each silhouette shape corresponds to the number of data points contained within that cluster, while its width indicates the sorted silhouette coefficients of the data points in that cluster; a wider silhouette implies better clustering. The vertical dashed lines in the diagrams denote the average silhouette coefficient for each set of clusters. The relative position of data points to this dashed line offers critical insights. Specifically, if many data points in a cluster don't reach or surpass this line, it indicates that those data points might be closely aligned with data from other clusters, thereby suggesting a lack of distinctness for that cluster. In contrast, when the majority of data points in a cluster largely surpass the dashed line, moving towards a coefficient of 1.0, it is indicative of a well-defined cluster. For instance, within the k = 6 silhouette diagram, there's a notable presence of data points in cluster 0 with negative silhouette coefficients, hinting that these might be more appropriately placed in another cluster. However, the k = 5 diagram presents a more coherent clustering pattern, as seen by most data points in its clusters surpassing the dashed line, hinting at more distinct groupings.

Figure 3.21 Sample Silhouette Diagrams for Various Values of k





3.3.3.4. Generalized Median String

The Levenshtein Median, also known as the string median or the sequence median, generalizes the concept of the median to sequences or strings. The Levenshtein Median, derived from a set of strings, represents a string (or sequence) that either exists within the initial collection or emerges as a completely new string not present in the original set. This characteristic aligns with Kohonen's (1985) generalized median, where the median does not necessarily have to be a member of the original set. Instead, the defining feature of the Levenshtein Median, consistent with Kohonen's perspective, is its unparalleled ability to minimize the total Levenshtein distance in relation to all other strings in the collection. Kohonen further elucidates that while a median can sometimes be found within the existing set, termed a 'set median', there are instances where it may manifest as an entirely new sequence. In this study, this concept was vital in identifying sequences that, on average, require the fewest modifications to match any other sequences within their cluster, offering profound insights into the core or 'center' of each cluster.

The Levenshtein Median was employed post-clustering to identify the most representative sequence within each established cluster of participants' engineering approaches, search strategies, and ideas. This process substantially aids in the interpretation of our clustering results.

Here is a general description of the algorithm to find the Levenshtein median:

1. Given a set of strings $S = \{s_1, s_2, ..., s_n\}$, the Levenshtein median, denoted as M, is the string that minimizes the sum of Levenshtein distances to all other strings in S.
2. Start with an initial estimate of M, which can be any string from the set S.
3. Iterate the following steps until convergence or a stopping condition is met:
    a. Calculate the Levenshtein distance between M and each string s_i in S.
    b. Update M as the string that minimizes the sum of distances obtained in step 3a.
4. Once the algorithm converges or the stopping condition is met, the resulting M is considered the Levenshtein median.

This method actively seeks to formulate a sequence, either from S or a novel one, that ensures the least aggregate distances from every other sequence in the collection. The direct calculation of the median string without employing any iterative optimization strategies has limitations in terms of computational efficiency and may not always yield the most accurate representation.

Using an iterative optimization strategy, such as the greedy algorithm, in Levenshtein Median is useful for several reasons. Firstly, it allows for a more refined search process, exploring different candidate strings and iteratively improving their total distance. This iterative approach helps in finding a better approximation of the true median and increases the likelihood of identifying the most representative string within the set.

The limitation of a simple Levenshtein Median, without iterative optimization strategies, is that it may produce suboptimal results. It might select a string that is not the most accurate representative of the set due to limited exploration of the solution space. Without iterative optimization, the method is restricted to a single computation, potentially missing out on better alternatives.

The greedy algorithm, employed in the Levenshtein.median() function of the python-Levenshtein package, addresses this limitation. By iteratively refining the candidate string, the greedy algorithm continuously improves the total distance to other strings in the set. It performs a step-by-step optimization, always selecting the locally optimal choice at each iteration. This approach ensures a more thorough exploration of the solution space, leading to a better approximation of the Levenshtein Median.

In this study, I adopted the greedy iterative algorithm provided by the python-Levenshtein package. This approach allowed efficiently computing an approximate generalized median string for each cluster. By using the greedy algorithm, I could identify the most representative sequence within each cluster, capturing the essence of participants' engineering approaches, search strategies, and ideas.

Chapter 4. Results

4.1. RQ 1. What progress in engineering approaches, search strategies, and integrating ideas do students make while using the electric circuit simulation in the control condition versus rule formation conditions?

4.1.1 Descriptive Statistics

Pretest scores for the concept of current, out of a total score of 5, indicated that the Control group had a mean score of $M = 1.45$, with a standard deviation of $SD = 0.86$. The Experiment 1 group (Decision Table and Rule Induction) exhibited a mean score of $M = 1.65$, with a standard deviation of $SD = 0.65$, while the Experiment 2 group (Rule Induction) displayed a mean score of $M = 1.15$, with a standard deviation of $SD = 0.96$.

Pretest results for the concept of voltage drop, out of a total score of 5, indicated that the Control group achieved a mean score of $M = 1.8$, with a standard deviation of $SD = 0.6$. The Experiment 1 group, which utilized the Decision Table and Rule Induction intervention, reported a mean score of $M = 1.5$, with a standard deviation of $SD = 0.87$. Meanwhile, participants in the Experiment 2 group, receiving the Rule Induction intervention, had a mean score of $M = 1.2$, with a standard deviation of $SD = 0.98$.

Table 4.1 illustrates the descriptive statistics for engineering approaches and Table 4.2 illustrates the descriptive statistics for search strategies. Table 4.1 presents descriptive statistics for the various engineering approaches adopted by the students across the three conditions: Control, Decision Table & Rule Induction, and Rule Induction. Similarly, Table 4.2 provides the descriptive statistics for the search strategies employed by the students. For each student in the study, I identified and counted the occurrences of each engineering approach and search strategy during the task. These counts were then used to calculate the mean (M) and standard deviation (SD) for each approach and strategy within each condition, providing a quantitative overview of the engineering strategies and search behaviors utilized by the students. This method of analysis allowed me to discern patterns and variations in the adoption of different engineering approaches and search strategies across the various conditions.

Table 4.1 Mean and Standard Deviation for Engineering Approaches

	Control N = 20 (1-20)		Decision Table & Rule Induction N = 20 (21-40)		Rule Induction N = 20 (41-60)	
	M	SD	M	SD	M	SD
1 bulbs equal	0.75	0.79	0.85	0.67	1.47	1.17
2 bulbs equal	0.7	0.86	1.25	0.91	0.89	0.94
1 vs. 2 bulbs different	0.05	0.22	0.5	0.76	0.53	0.84
Battery voltage equal	1.05	1.05	2.55	1	2.11	1.05
Battery voltage different	0.45	0.51	0.05	0.22	0.79	0.71
Resistance change: 1 or 2 bulbs	0.45	0.60	0.55	0.69	0.47	0.84
Resistance change: 2 bulbs	0.05	0.22	0.45	0.83	0.21	0.42
Resistance equal: 2 bulbs different value	0.4	0.75	0.7	0.92	0.42	0.61
Resistance equal: 2 or 1 vs. 2 bulbs same value	0.25	0.44	0.2	0.41	0.47	0.77
Resistance equal: 1 bulb	0.35	0.49	0.7	0.57	1.26	1.05
Resistance equal: 2 bulbs different value & position	0	0	0.1	0.31	0.11	0.32
Ammeter equal: after/before 1 or 2 bulb	0.85	0.59	1.15	0.81	1.74	1.10
Ammeter different: before vs. after 1 or 2 bulbs	0.45	0.76	0.95	0.69	1.16	0.69
Ammeter different: on high/low of 2 bulb	0.2	0.52	0.5	0.51	0	0
Voltmeter equal: 1 bulb or 2 bulbs or battery	0.95	0.76	2.1	1.17	2.47	1.17
Voltmeter equal: 2 bulbs	0	0	0.05	0.22	0.16	0.50
Voltmeter different: high vs. low resistance 2 bulbs	0.3	0.57	0.5	0.69	0.47	0.61

Voltmeter different: high vs. low resistance 2 bulbs vs. battery	0.15	0.37	0.1	0.31	0.11	0.46
No action	0.25	0.55	0	0	0	0

Table 4.2 Mean and Standard Deviation for Search Strategies

	Control N = 20 (1-20)		**Decision Table & Rule Induction N = 20 (21-40)**		**Rule Induction N = 20 (41-60)**	
	M	SD	M	SD	M	SD
No action	1.75	1.68	0.55	0.51	0.79	0.42
No new comparative trial	0.6	0.5	0	0	0	0
Prediction Current: same rule	0.8	0.52	1.25	1.02	2.05	1.47
Prediction Voltage Drop: same rule	0.95	0.89	1.65	0.99	1.68	1.34
Confidence Current: verify prediction	0.95	0.6	1.65	1.04	2.21	1.32
Confidence Voltage Drop: verify prediction	0.8	0.62	1.9	1.21	1.89	1.24
Rule Current: Confirming Redundancy	0.75	0.55	0.3	0.66	0.47	0.61
Rule Voltage Drop: Confirming Redundancy	0.55	0.6	0.2	0.52	0.16	0.37
Post-test: Current non-normative	0.65	0.49	0.1	0.31	0.11	0.32
Post-test: Voltage Drop non-normative	0.65	0.49	0.05	0.22	0.05	0.23
New comparative trial	0.4	0.5	1	0	1	0
Prediction Current: Fill up gaps	0.5	0.89	1.35	0.99	0.84	0.6
Prediction Voltage Drop: Fill up gaps	0.35	0.67	0.95	0.83	1.21	0.71
Confidence Current: falsify prediction	0.35	0.59	0.95	0.76	0.68	0.48
Confidence Voltage Drop: falsify prediction	0.3	0.73	0.7	0.66	1	0.58

Identify problem: goal stated	0.15	0.37	0.45	0.51	0.21	0.42
Rule Current: Simultaneous scanning	0.1	0.45	0.85	0.81	1.42	1.07
Rule Voltage Drop: Simultaneous scanning	0.25	0.64	1.05	0.69	0.74	0.45
Rule Current: Successive scanning	0.1	0.31	0.85	0.49	0.84	0.37
Rule Voltage Drop: Successive scanning	0.1	0.31	0.8	0.62	1.21	0.79
Rule Current: Focus gambling	0.25	0.44	0.15	0.37	0.11	0.32
Rule Voltage Drop: Focus gambling	0.25	0.44	0.3	0.47	0.26	0.45
Rule Current: Conservative Focusing	0.2	0.52	0.45	0.69	0.05	0.23
Rule Voltage Drop: Conservative Focusing	0.1	0.31	0.2	0.41	0.16	0.37
Post-test: Current partial	0.15	0.37	0.35	0.49	0.53	0.51
Post-test: Voltage Drop partial	0.15	0.37	0.45	0.51	0.58	0.51
Post-test: Current 1 Valid link	0.05	0.22	0.1	0.31	0.32	0.48
Post-test: Voltage Drop 1 Valid link	0.1	0.31	0.2	0.41	0	0
Post-test: Current 2 Valid links	0.15	0.37	0.45	0.51	0.05	0.23
Post-test: Voltage Drop 2 Valid links	0.1	0.31	0.3	0.47	0.37	0.5

Figure 4.1 and Figure 4.2 illustrate the frequencies of each variable among the Control group, Experimental 1 group (using the decision table and induction rule tool), and Experimental 2 group (using the induction rule tool).

Figure 4.1 Count of Comparative Trials for Engineering Approaches among Control and Experimental Groups

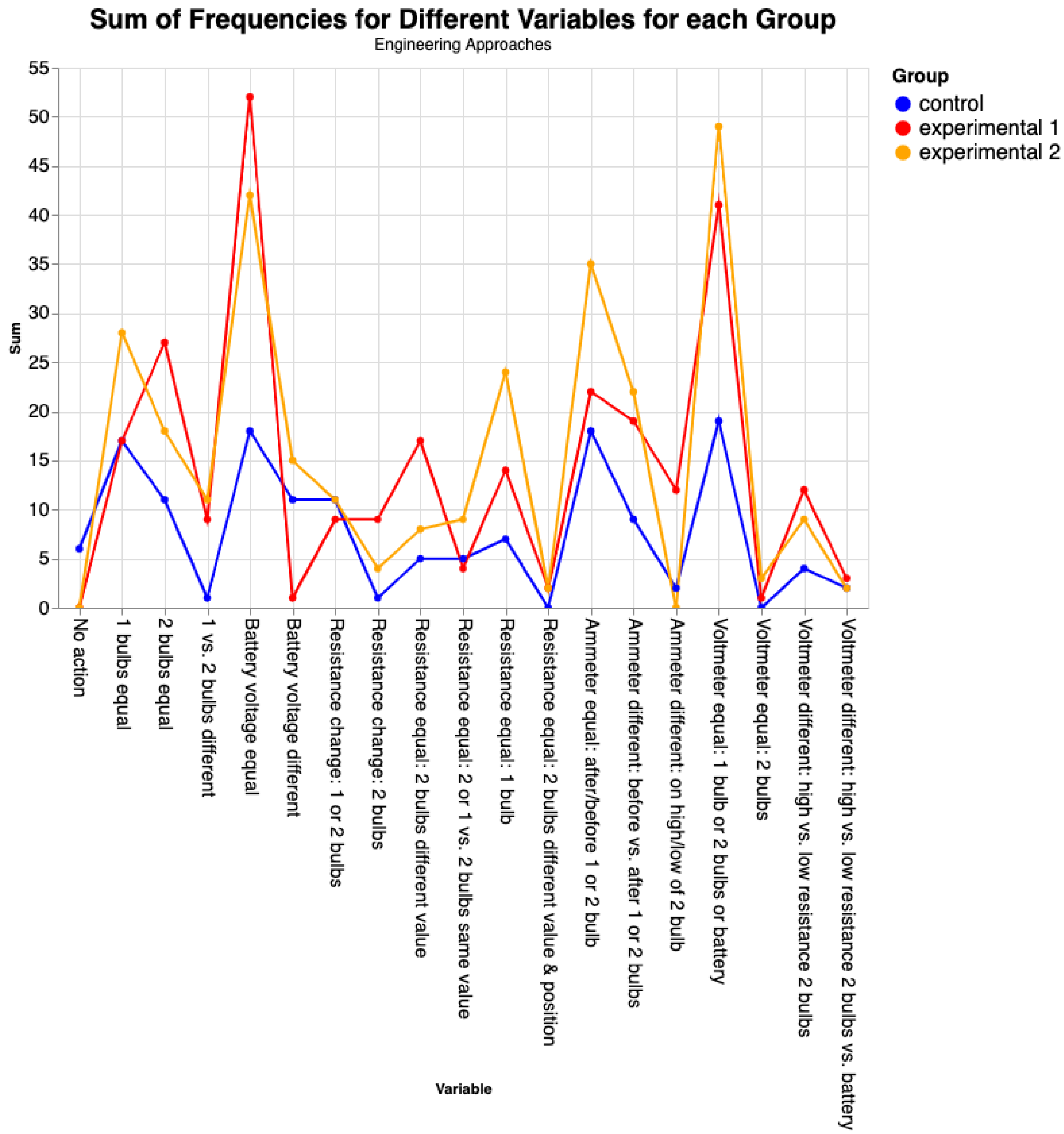
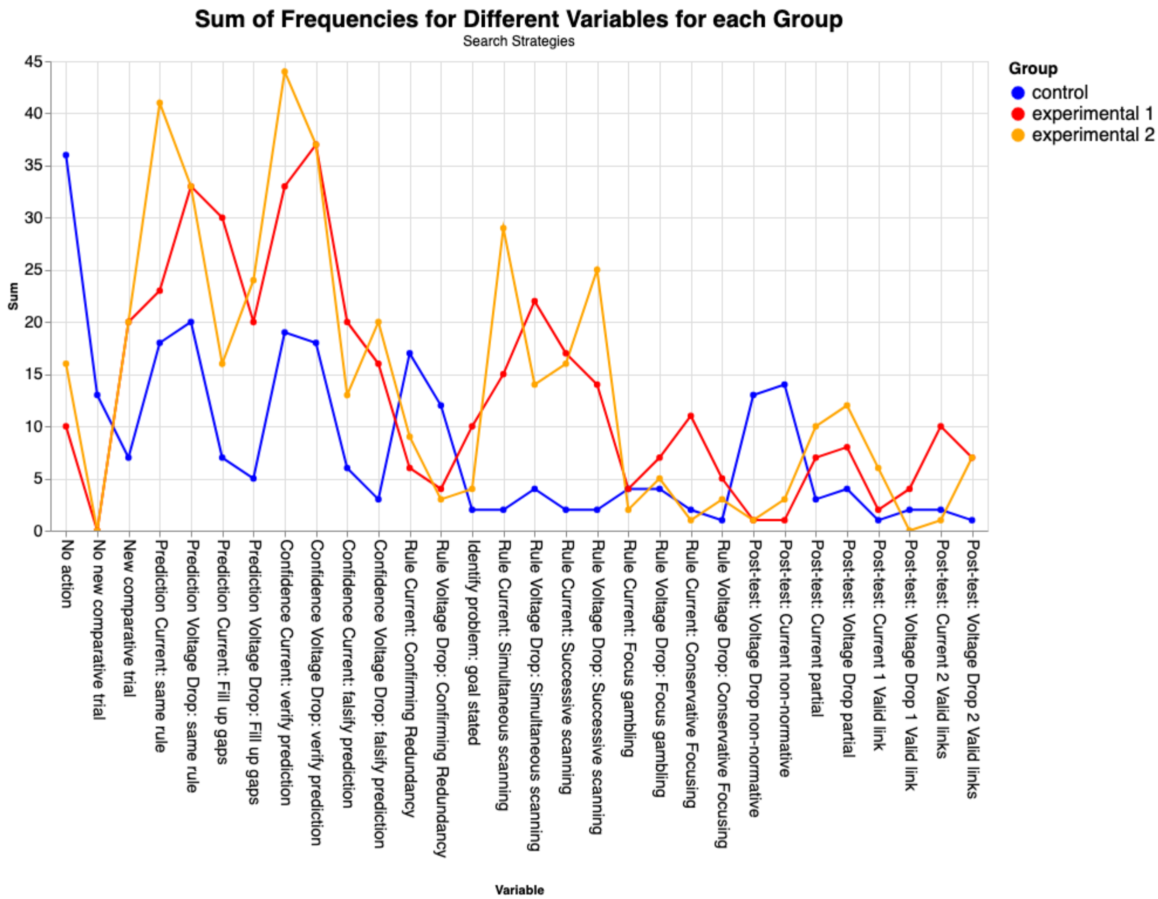


Figure 4.2 Count of Comparative Trials for Search Strategies among Control and Experimental Groups



4.1.2. Dunn’s Test and Spearman Correlation Coefficient

This investigation centered around a set of thirty search strategy variables, crucial to the hypothesis that, by utilizing a decision table and a dynamic rule induction tool, individuals could strategically plan actions based on their comparative selection of conditions. The thirty search strategy variables are:

- No action
- New comparative trial
- No new comparative trial
- Prediction Current: same rule
- Prediction Voltage Drop: same rule
- Prediction Current: Fill up gaps
- Prediction Voltage Drop: Fill up gaps
- Confidence Current: verify prediction
- Confidence Voltage Drop: verify prediction

- Confidence Current: falsify prediction
- Confidence Voltage Drop: falsify prediction
- Identify problem: goal stated
- Rule Current: Confirming Redundancy
- Rule Voltage Drop: Confirming Redundancy
- Rule Current: Simultaneous scanning
- Rule Voltage Drop: Simultaneous scanning
- Rule Current: Successive scanning
- Rule Voltage Drop: Successive scanning
- Rule Current: Focus gambling
- Rule Voltage Drop: Focus gambling
- Rule Current: Conservative Focusing
- Rule Voltage Drop: Conservative Focusing
- Post-test: Current non-normative
- Post-test: Voltage Drop non-normative
- Post-test: Current partial
- Post-test: Voltage Drop partial
- Post-test: Current 1 Valid link
- Post-test: Voltage Drop 1 Valid link
- Post-test: Current 2 Valid links
- Post-test: Voltage Drop 2 Valid links

Upon examining these variables using the Shapiro-Wilk test, none adhered to the assumption of normality, a key prerequisite for ANOVA. The non-normal distribution of data in this study could be attributed to the intrinsic characteristics of the search strategy variables, which were predominantly categorical or ordinal in nature. These variables often recorded scores at lower ends of the spectrum, predominantly 0 and 1, with some variables extending to values up to 5 but not higher. Given these variables' limited range and the tendency for scores to aggregate at the lower end of the scale, the resulting distributions deviate from the normality typically expected in continuous data. This aggregation effect, particularly pronounced due to the variables' discrete and constrained scoring system, likely contributed to the observed departure from normal distribution patterns. Since the assumptions for performing ANOVA are not met, it is appropriate to use a non-parametric test. Consequently, Dunn's test was selected for this analysis. It is a post-hoc test that can be used to compare the mean ranks of three or more groups when the assumption of normality is violated or when the data is ordinal. It is a good alternative to ANOVA in such situations. To mitigate the potential for elevated Type I error due to numerous comparisons, I employed the Benjamini-Hochberg procedure, modifying the significance level for every statistical analysis in this section. As a result, the *p*-values presented have been adjusted considering these multiple tests.

Spearman correlation coefficient (calculated by the spearman r function from scipy.stats) is a measure of rank correlation. This means that it assesses monotonic relationships between two variables using the ranks of the values rather than the values themselves. This is particularly appropriate for data where the relationship might not be linear, or when the data is not normally distributed. I calculated the Spearman correlation for each variable separately between two groups (Control and Experimental). I looked at whether the ranks of the values for a given variable are similarly ordered in both the control group and the experimental group.

4.1.2.1. Search Strategies for Experimental 1 (Decision Table and Induction Rule Tool) vs. Control

Dunn's test was conducted to compare the differences between the Control and Experimental group 1 (see Figure 4.3, Table 4.3). The correlation between the Control and Experimental 1 groups for search strategy variables was computed using Spearman's rho, and the effect size was rank-biserial correlation.

Statistically detectable differences were found in the 'No action' condition, $z = 3.25$, $p < .001$, with a negative correlation ($r = -.26$), and a large effect size ($rs = 1.00$). For the 'No new comparative trial' condition, results were significant, $z = 3.52$, $p < .001$, with a positive correlation (rs = N/A), and a large effect size ($r = 1.00$). In the 'New comparative trial' condition, results were also statistically detectable, $z = -3.52$, $p < .001$, with a positive correlation (rs = N/A), and a large effect size ($r = 1.00$).

For the prediction variables, significant differences were found for 'Prediction Voltage Drop: same rule' ($z = -2.14$, $p = .02$, $rs = .13$, $r = 1.00$), 'Prediction Current: Fill up gaps' ($z = -3.65$, $p < .001$, $rs = .01$, $r = 1.00$), 'Prediction Voltage Drop: Fill up gaps' ($z = -2.80$, $p < .001$, $rs = .17$, $r = 1.00$). However, for the 'Prediction Current: same rule' condition, there was no significant difference, z = -0.46, p = .60, with a negative correlation (rs = .10), and a large effect size (r = 1.00).

For the confidence variables, the results of 'Confidence Current: verify prediction' (z = -2.11, p = .02, rs = .14, r = 1.00), 'Confidence Voltage Drop: verify prediction' (z = -2.52, p = .01, rs = .11, r = 1.00), 'Confidence Current: falsify prediction' (z = -2.84, p < .001, rs = .27, r = 1.00), and 'Confidence Voltage Drop: falsify prediction' (z = -2.56, p < .001, rs = .25, r = 1.00) conditions were statistically detectable.

For problem identification and rule formation variables, statistically detectable differences were found for 'Identify problem: goal stated' (z = -2.16, p = .01, rs = .00, r = 1.00), 'Rule Current: Confirming Redundancy' (z = 2.68, p < .001, rs = .10, r = 1.00), and 'Rule Voltage Drop: Confirming Redundancy' (z = 2.06, p = .01, rs = .04, r = 1.00). There were also significant differences observed for 'Rule Current: Conservative Focusing' (z = -1.70, p = .02, rs = .15, r = 1.00). However, the 'Rule Current: Focus gambling' condition showed no significant difference (z = 0, p = 1.00, rs = .25, r = 1.00), and 'Rule Voltage Drop: Focus gambling' condition was also not significant (z = -0.81, p = .29, rs = .37, r = 1.00). Similarly, the 'Rule Voltage Drop: Conservative Focusing' condition showed no significant difference (z = -1.08, p = .08, rs = .13, r = 1.00).

For the post-test variables, the 'Post-test: Voltage Drop non-normative' and 'Post-test: Current non-normative' conditions, results were significantly different, z = 3.25, p < .001, rs = .17, r = 1.00, and z = 3.52, p < .001, rs = .35, r = 1.00, respectively. Similarly, significant differences were found for 'Post-test: Current 2 Valid links' (z = -2.16, p = .01, rs = .00, r = 1.00), and 'Post-test: Voltage Drop 2 Valid links' (z = -1.62, p = .02, rs = .17, r = 1.00). However, for the 'Post-test: Current partial' condition, no significant difference was observed (z = -1.08, p = .15, rs = .28, r = 1.00), and 'Post-test: Voltage Drop partial' also showed no significant difference (z = -1.08, p = .17, rs = .36, r = 1.00). Furthermore, no significant differences were found for 'Post-test: Current 1 Valid link' (z = -0.27, p = .55, rs = .08, r = 1.00), and 'Post-test: Voltage Drop 1 Valid link' (z = -0.54, p = .38, rs = .67, r = 1.00).

Figure 4.3 Dunn's Test Comparisons: Control vs. Experimental 1 (Decision Table and Rule Formation)

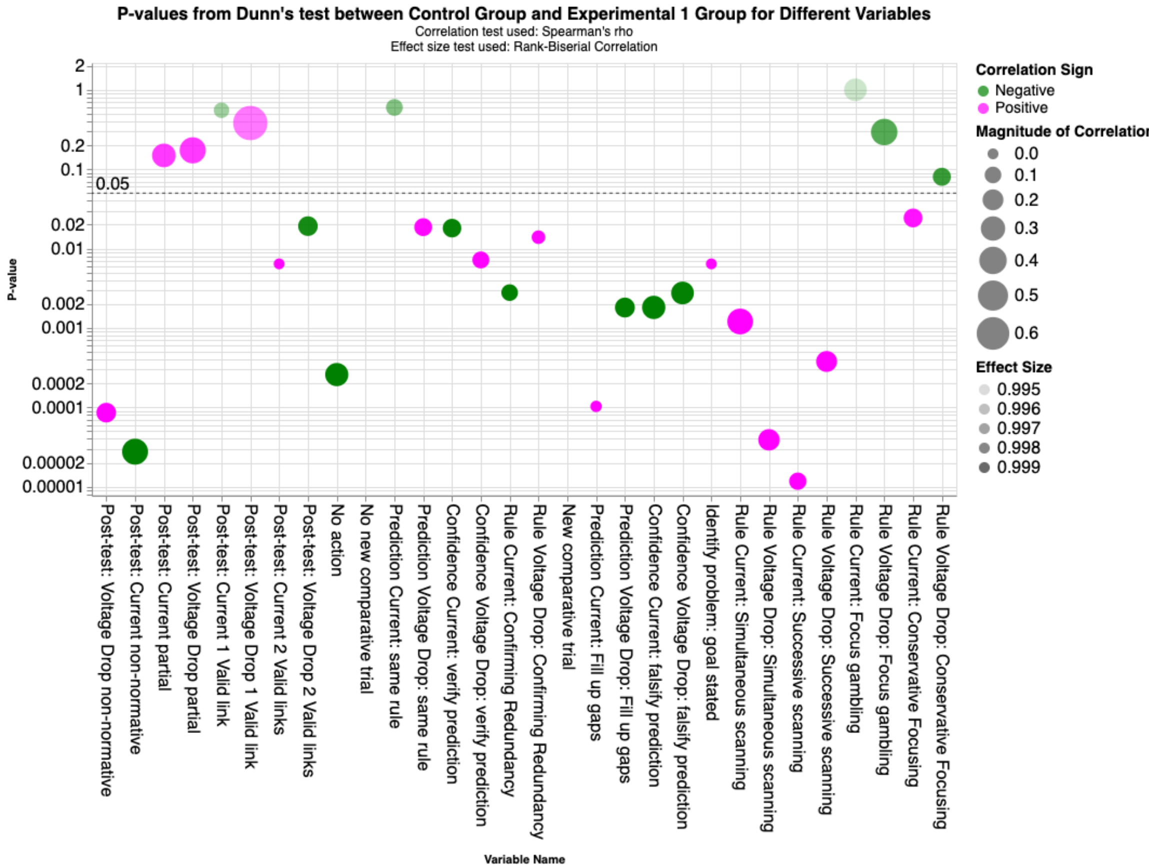


Table 4.3 Significant P-values Comparing Control vs. Experimental 1 (Decision Table and Rule Formation) for Each Variable

Variable	P-value	Effect Size	Z-score	Correlation	Correlation Sign
No action	0	1	3.25	0.26	Negative
No new comparative trial	0	1	3.52	N/A	Positive
New comparative trial	0	1	-3.52	N/A	Positive
Prediction Voltage Drop: same rule	0.02	1	-2.14	0.13	Positive
Prediction Current: Fill up gaps	0	1	-3.65	0.01	Positive
Prediction Voltage Drop: Fill up gaps	0	1	-2.8	0.17	Negative
Confidence Current: verify prediction	0.02	1	-2.11	0.14	Negative
Confidence Voltage Drop: verify prediction	0.01	1	-2.52	0.11	Positive
Confidence Current: falsify prediction	0	1	-2.84	0.27	Negative

Confidence Voltage Drop: falsify prediction	0	1	-2.56	0.25	Negative
Identify problem: goal stated	0.01	1	-2.16	0	Positive
Rule Current: Confirming Redundancy	0	1	2.68	0.1	Negative
Rule Voltage Drop: Confirming Redundancy	0.01	1	2.06	0.04	Positive
Rule Current: Simultaneous scanning	0	1	-2.61	0.34	Positive
Rule Voltage Drop: Simultaneous scanning	0	1	-3.73	0.21	Positive
Rule Current: Successive scanning	0	1	-3.81	0.11	Positive
Rule Voltage Drop: Successive scanning	0	1	-3	0.2	Positive
Rule Current: Conservative Focusing	0.02	1	-1.7	0.15	Positive
Post-test: Voltage Drop non-normative	0	1	3.25	0.17	Positive
Post-test: Current non-normative	0	1	3.52	0.35	Negative
Post-test: Current 2 Valid links	0.01	1	-2.16	0	Positive
Post-test: Voltage Drop 2 Valid links	0.02	1	-1.62	0.17	Negative

4.1.2.2. Search Strategies for Experimental 2 (Induction Rule Tool) vs. Control

Dunn's test was conducted to compare the differences between the Control and Experimental group 2 (see Figure 4.4 & Table 4.4). The correlation test used was Spearman's rho, and the effect size test used was rank-biserial correlation.

The Dunn's test showed a significant positive correlation in the "No action" variable ($Z = 2.11$, $p = .01$, $rs = 0.43$, $r = 1$), while the "No new comparative trial" variable also revealed a significant positive correlation ($Z = 3.52$, $p < .001$, $r = 1$). Similarly, the "New comparative trial" ($Z = -3.52$, $p < .001$, $r = 1$) showed significant positive correlation.

For the prediction variables, "Prediction Current: same rule" variables showed a statistically detectable positive correlation ($Z = -2.62$, $p < .001$, $rs = 0.3$, $r = 1$). However, the "Prediction Voltage Drop: same rule" variable, while exhibiting a positive correlation, did not reach a statistically detectable difference ($Z = -1.73$, $p = .06$, $rs = 0.38$, $r = 1$). On the other hand, the "Prediction Current: Fill up gaps" variable revealed a significant negative correlation ($Z = -2.19$, $p = .01$, $rs = 0.43$, $r = 1$). Nevertheless, significant positive correlations were observed in the "Prediction Voltage Drop: Fill up gaps" ($Z = -3.67$, $p < .001$, $rs = 0.09$, $r = 1$).

For the confidence variables, the "Confidence Current: falsify prediction" revealed a significant negative correlation ($Z = -1.99$, $p = .02$, $rs = 0.1$, $r = 1$). Nevertheless, significant positive correlations were observed in the "Confidence Current: verify prediction", "Confidence Voltage Drop: verify prediction", and "Confidence Voltage Drop: falsify prediction" variables ($Z = -3.18$, $p < .001$, $rs = 0.59$, $r = 1$; $Z = -2.54$, $p = .01$, $rs = 0.12$, $r = 1$; and $Z = -3.91$, $p < .001$, $rs = 0.26$, $r = 1$, respectively).

For the problem identification and rule formation variables, the "Identify problem: goal stated" variable showed a positive correlation but was not statistically detectable ($Z = -0.54$, $p = .38$, $rs = 0.25$, $r = 1$), while the "Rule Current: Confirming Redundancy" variable demonstrated a significant positive correlation ($Z = 1.91$, $p = .03$, $rs = 0.07$, $r = 1$). The "Rule Voltage Drop: Confirming Redundancy" variable revealed a significant negative correlation ($Z = 2.2$, $p = .01$, $rs = 0.19$, $r = 1$). Positive correlations were also found in the "Rule Current: Simultaneous scanning", "Rule Voltage Drop: Simultaneous scanning", "Rule Current: Successive scanning", and "Rule Voltage Drop: Successive scanning" variables, all of which were statistically detectable ($Z = -4.25$, $p < .001$, $rs = 0.19$, $r = 1$; $Z = -2.87$, $p < .001$, $rs = 0.22$, $r = 1$; $Z = -3.79$, $p < .001$, $rs = 0.17$, $r = 1$; and $Z = -4.25$, $p < .001$, $rs = 0.13$, $r = 1$, respectively). In contrast, the "Rule Current: Focus gambling" variable showed a negative correlation ($Z = 0.54$, $p = .38$, $r = 0.17$, $r = 1$), while the "Rule Voltage Drop: Focus gambling" variable demonstrated a positive correlation ($Z = -0.27$, $p = .71$, $rs = 0$, $r = 1$). The "Rule Current: Conservative Focusing" variable revealed a negative correlation ($Z = 0.27$, $p = .55$, $rs = 0.08$, $r = 1$), and the "Rule Voltage Drop: Conservative Focusing" variable showed a positive correlation ($Z = -0.54$, $p = .3$, $rs = 0.55$, $r = 1$).

For the post-test variables, the 'Post-test: Current non-normative' condition, a positive correlation was found ($r = .28$, $z = 2.98$, $p < .001$, $rs = 1.00$). In the 'Post-test: Current partial' condition, results were statistically detectable ($r = .14$, $z = -1.89$, $p = .02$, $rs = 1.00$) and showed a positive correlation. Similarly, for the 'Post-test: Voltage Drop partial' condition, a significant difference was found ($r = .10$, $z = -2.16$, $p = .01$, $rs = 1.00$), but the correlation was negative. For the 'Post-test: Current 1 Valid link' condition, the results showed a negative correlation ($r = .15$, $z = -1.35$, $p = .04$, $rs = 1.00$), signifying a detectable difference. However, the 'Post-test: Voltage Drop 1 Valid link' condition was not significantly different (z = 0.54, p = .15, rs = 1.00), with a positive correlation being unable to be calculated. Further, for the 'Post-test: Current 2 Valid links' condition, there was no significant difference (r = .69, z = 0.27, p = .55, rs = 1.00), with a positive correlation. Conversely, the 'Post-test: Voltage Drop 2 Valid links' condition was significantly different (r = .31, z = -1.62, p = .02, rs = 1.00), with a positive correlation.

Figure 4.4 Dunn’s Test Comparisons: Control vs. Experimental 2 (Rule Formation)

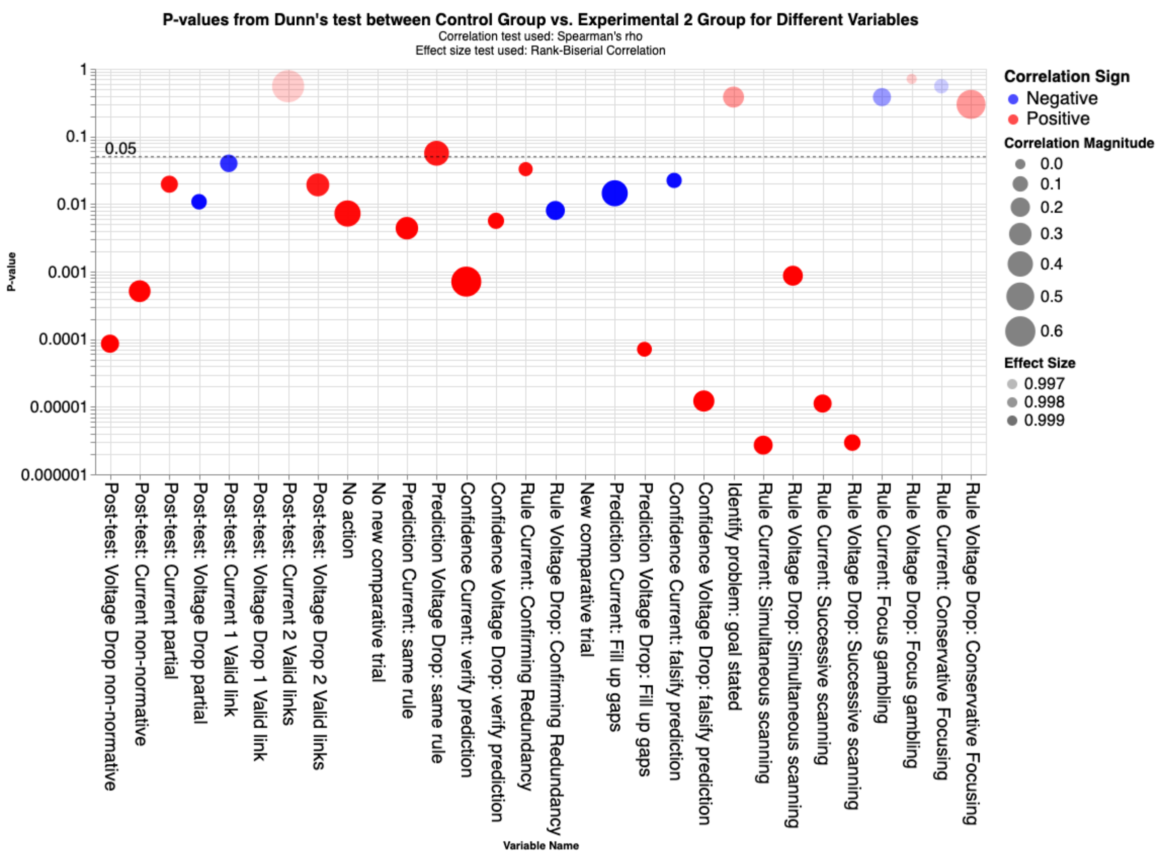


Table 4.4 Significant P-values Comparing Control vs. Experimental 2 (Rule Formation) for Each Variable

Variable	P-value	Effect Size	Z-score	Correlation	Correlation Sign
No action	0.01	1	2.11	0.43	Positive
No new comparative trial	0	1	3.52		Positive
New comparative trial	0	1	-3.52		Positive

Prediction Current: same rule	0	1	-2.62	0.3	Positive
Prediction Current: Fill up gaps	0.01	1	-2.19	0.43	Negative
Prediction Voltage Drop: Fill up gaps	0	1	-3.67	0.09	Positive
Confidence Current: verify prediction	0	1	-3.18	0.59	Positive
Confidence Voltage Drop: verify prediction	0.01	1	-2.54	0.12	Positive
Confidence Current: falsify prediction	0.02	1	-1.99	0.1	Negative
Confidence Voltage Drop: falsify prediction	0	1	-3.91	0.26	Positive
Rule Current: Confirming Redundancy	0.03	1	1.91	0.07	Positive
Rule Voltage Drop: Confirming Redundancy	0.01	1	2.2	0.19	Negative
Rule Current: Simultaneous scanning	0	1	-4.25	0.19	Positive
Rule Voltage Drop: Simultaneous scanning	0	1	-2.87	0.22	Positive
Rule Current: Successive scanning	0	1	-3.79	0.17	Positive
Rule Voltage Drop: Successive scanning	0	1	-4.25	0.13	Positive
Post-test: Voltage Drop non-normative	0	1	3.25	0.17	Positive
Post-test: Current non-normative	0	1	2.98	0.28	Positive
Post-test: Current partial	0.02	1	-1.89	0.14	Positive
Post-test: Voltage Drop partial	0.01	1	-2.16	0.1	Negative
Post-test: Current 1 Valid link	0.04	1	-1.35	0.15	Negative
Post-test: Voltage Drop 2 Valid links	0.02	1	-1.62	0.31	Positive

4.1.2.3. Search Strategies for Experimental 1 vs. Experimental 2

Dunn's test was conducted to compare the differences between the Experimental group 1 and Experimental group 2 (see Figure 4.5 & Table 4.5). The correlation test used was Spearman's rho, and the effect size test used was rank-biserial correlation.

In the Dunn's test analysis comparing Experimental 1 and Experimental 2 groups, the variable "No action" showed a significant positive correlation ($Z = -1.62$, $p = .05$, $r = 1$, $rs = 0$). There was no available p-value or correlation for the "No new comparative trial" and "New comparative trial" variables; however, they demonstrated a large effect size ($r = 1$).

For prediction variables, a significant negative correlation was found in "Prediction Current: same rule" ($Z = -2.03$, $p = .03$, $r = 1$, $rs = 0.17$), while the variables "Prediction Voltage Drop: same rule" and "Prediction Current: Fill up gaps" both showed a significant positive correlation ($Z = 0.34$, $p = .72$, $r = 1$, $rs = 0.1$; and $Z = 2.19$, $p = .02$, $r = 1$, $rs = 0.01$, respectively). The "Prediction Voltage Drop: Fill up gaps" variable showed a positive correlation but was not statistically detectable ($Z = -0.89$, $p = .34$, $r = 1$, $rs = 0.14$)

For the confidence variables, "Confidence Current: verify prediction" and "Confidence Voltage Drop: verify prediction" revealed a negative correlation, but neither were statistically detectable ($Z = -1.41$, $p = .14$, $r = 1$, $rs = 0.36$; and $Z = 0.05$, $p = .95$, $r = 1$, $rs = 0$, respectively). For "Confidence Current: falsify prediction," a positive correlation was shown, but it was not statistically detectable ($Z = 1.42$, $p = .11$, $r = 1$, $rs = 0.3$). In contrast, "Confidence Voltage Drop: falsify prediction" demonstrated a negative correlation without a statistically detectable difference ($Z = -1.11$, $p = .21$, $r = 1$, $rs = 0.05$).

Further in the analysis, the variables "Rule Current: Confirming Redundancy" and "Rule Voltage Drop: Confirming Redundancy" revealed a negative correlation, but neither reached a statistically detectable difference ($Z = -0.92$, $p = .25$, $r = 1$, $rs = 0.19$; and $Z = 0.04$, $p = .95$, $r = 1$, $rs = 0.18$, respectively). A significant positive correlation was observed for "Identify problem: goal stated" ($Z = 1.62$, $p = .05$, $r = 1$, $rs = 0$). For the variables "Rule Current: Simultaneous scanning" and "Rule Voltage Drop: Simultaneous scanning," a significant negative correlation and a significant positive correlation were found respectively ($Z = -2.06$, $p = .03$, $r = 1$, $rs = 0.11$; and $Z = 1.76$, $p = .04$, $r = 1$, $rs = 0.11$, respectively). The variables "Rule Current: Successive scanning," "Rule Voltage Drop: Successive scanning," "Rule Current: Focus gambling," and "Rule Voltage Drop: Focus gambling" all showed positive correlations, but none were statistically detectable ($Z = 0.22$, $p = .77$, $r = 1$, $rs = 0.1$; $Z = -2.1$, $p = .02$, $r = 1$, $rs = 0.27$; $Z = 0.54$, $p = .38$, $r = 1$, $rs = 0.25$; and $Z = 0.54$, $p = .5$, $r = 1$, $rs = 0.06$, respectively). A significant positive correlation was found for "Rule Current: Conservative Focusing" ($Z = 1.93$, $p = .01$, $r = 1$, $rs = 0.2$), while the "Rule Voltage Drop: Conservative Focusing" variable showed a negative correlation but did not reach statistically detectable difference ($Z = 0.54$, $p = .44$, $r = 1$, $rs = 0.24$).

For the post-test variables, the "Post-test: Voltage Drop non-normative" variable, a negative correlation was observed, but it was not statistically detectable (Z = 0, p = 1, r = 1, rs = 0.05). The "Post-test: Current non-normative" variable demonstrated a negative correlation as well, but also without a statistically detectable difference (Z = -0.54, p = .3, r = 1, rs = 0.1). The variables "Post-test: Current partial" and "Post-test: Voltage Drop partial" showed negative correlations, but neither reached a statistically detectable difference (Z = -0.81, p = .34, r = 1, rs = 0.31; and Z = -1.08, p = .21, r = 1, rs = 0.38, respectively). For "Post-test: Current 1 Valid link", there was a negative correlation but it was not statistically detectable (Z = -1.08, p = .12, r = 1, rs = 0.22). Conversely, "Post-test: Voltage Drop 1 Valid link" showed a detectable difference with a positive correlation (Z = 1.08, p = .04, r = 1). Furthermore, “Post-test: Current 2 Valid links” variable showed a detectable difference with a negative correlation (Z = 2.43, p < 0.01, r = 1, rs = 0.23).

Figure 4.5 Dunn’s Test Comparisons: Experimental 1 vs. Experimental 2

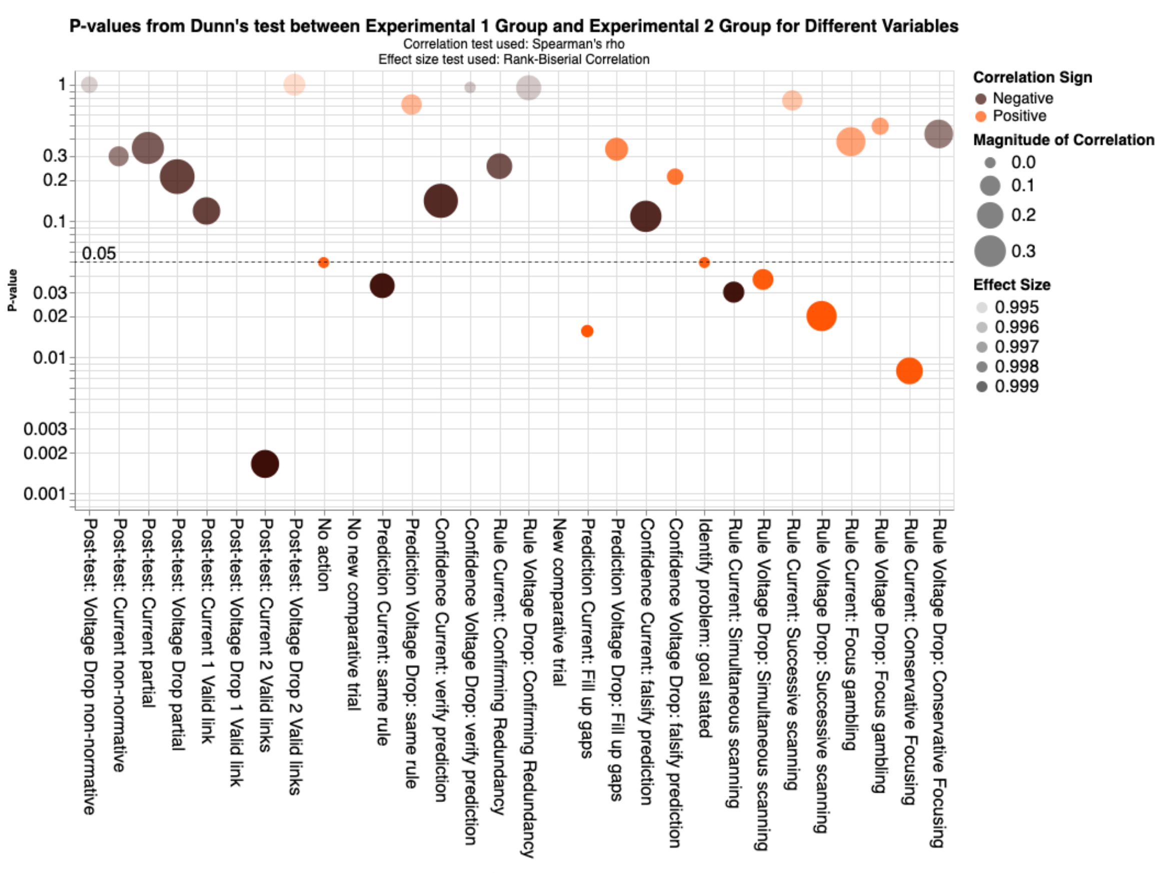


Table 4.5 Significant P-values for Experimental 1 vs. Experimental 2

Variable	P-value	Effect Size	Z-score	Correlation	Correlation Sign
No action	0.05	1	-1.62	0	Positive
Prediction Current: same rule	0.03	1	-2.03	0.17	Negative
Prediction Current: Fill up gaps	0.02	1	2.19	0.01	Positive
Identify problem: goal stated	0.05	1	1.62	0	Positive
Rule Current: Simultaneous scanning	0.03	1	-2.06	0.11	Negative
Rule Voltage Drop: Simultaneous scanning	0.04	1	1.76	0.11	Positive
Rule Voltage Drop: Successive scanning	0.02	1	-2.1	0.27	Positive
Rule Current: Conservative Focusing	0.01	1	1.93	0.2	Positive
Post-test: Voltage Drop 1 Valid link	0.04	1	1.08	N/A	Positive
Post-test: Current 2 Valid links	0	1	2.43	0.23	Negative

4.1.2.4. Comparing the Three Groups

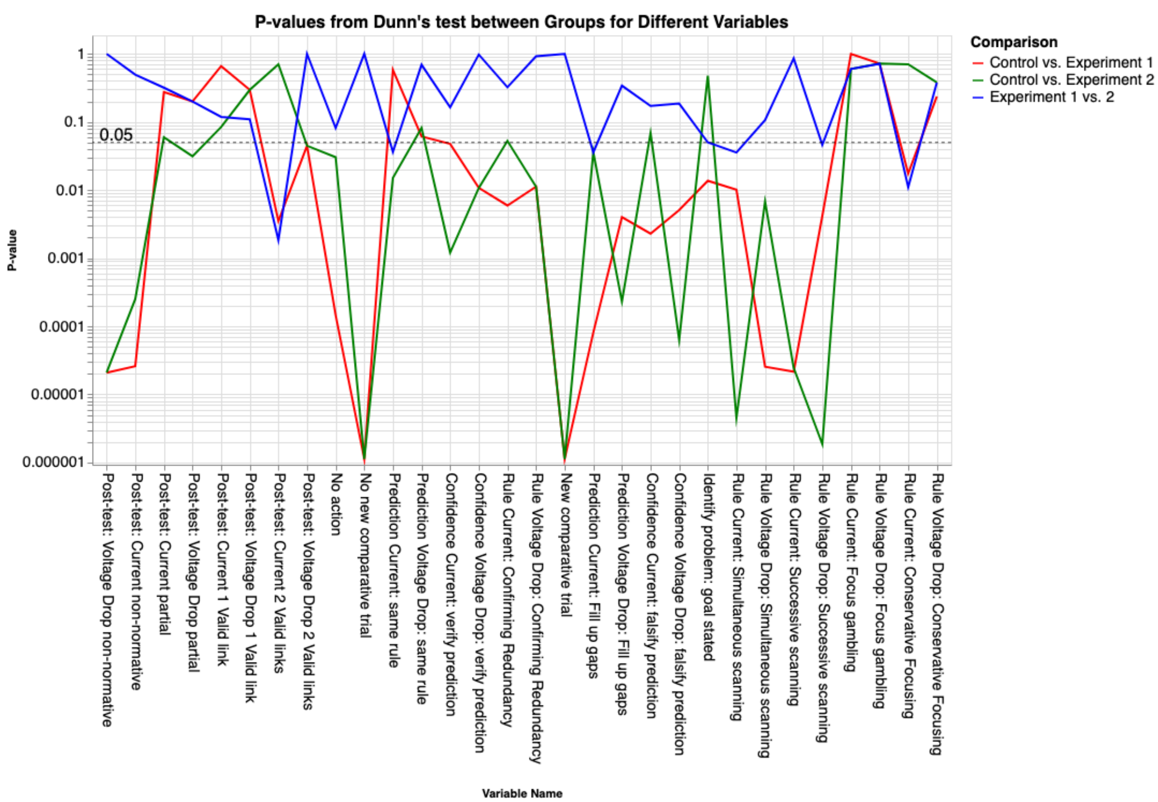
Figure 4.6 shows that the following variables are detectably different between Control vs. Experimental 1 but are not detectably different between Control vs. Experimental 2:

- Prediction Voltage Drop: same rule
- Confidence Current: verify prediction
- Confidence Voltage Drop: verify prediction
- Confidence Current: falsify prediction
- Confidence Voltage Drop: falsify prediction
- Identify problem: goal stated
- Rule Current: Confirming Redundancy
- Rule Voltage Drop: Confirming Redundancy
- Rule Current: Conservative Focusing
- Post-test: Voltage Drop non-normative
- Post-test: Current non-normative
- Post-test: Current 2 Valid links

Furthermore, the variables that are detectably different in both Control vs. Experimental group 1 and Experimental group 1 vs. Experimental group 2 are the following (Figure 4.6):

- No action
- Prediction Current: same rule
- Prediction Current: Fill up gaps
- Identify problem: goal stated
- Rule Current: Simultaneous scanning
- Rule Voltage Drop: Simultaneous scanning
- Rule Voltage Drop: Successive scanning
- Rule Current: Conservative Focusing
- Post-test: Current 2 Valid links

Figure 4.6 Dunn's Test Comparisons: Control vs. Experimental 1, Control vs. Experimental 2, and Experimental 1 vs. Experimental 2



4.2. RQ 2. What distinct learning paths do students take as they engage with the electric circuit simulation in the control condition versus intervention conditions?

4.2.1. Sequence Analysis between Control and Experimental Conditions

Figure 4.7 shows the Silhouette scores against the values of clusters. Among different clusters, k = 5 has the highest silhouette score average while k = 2 has the lowest silhouette score average (Figure 4.7). Among different clusters k = 5 has the least negative values. Figure 4.8 is the graphical Silhouette Coefficient clustering that shows that k = 5 has no negative values. Thus, I selected k = 5 as the optimal number of clusters.

After testing different numbers of clusters, five clusters showed more inter-cluster similarity and intra-cluster dissimilarity (Figure 4.9).

Figure 4.7 Silhouette Scores for Clusters for all Groups (N = 60)

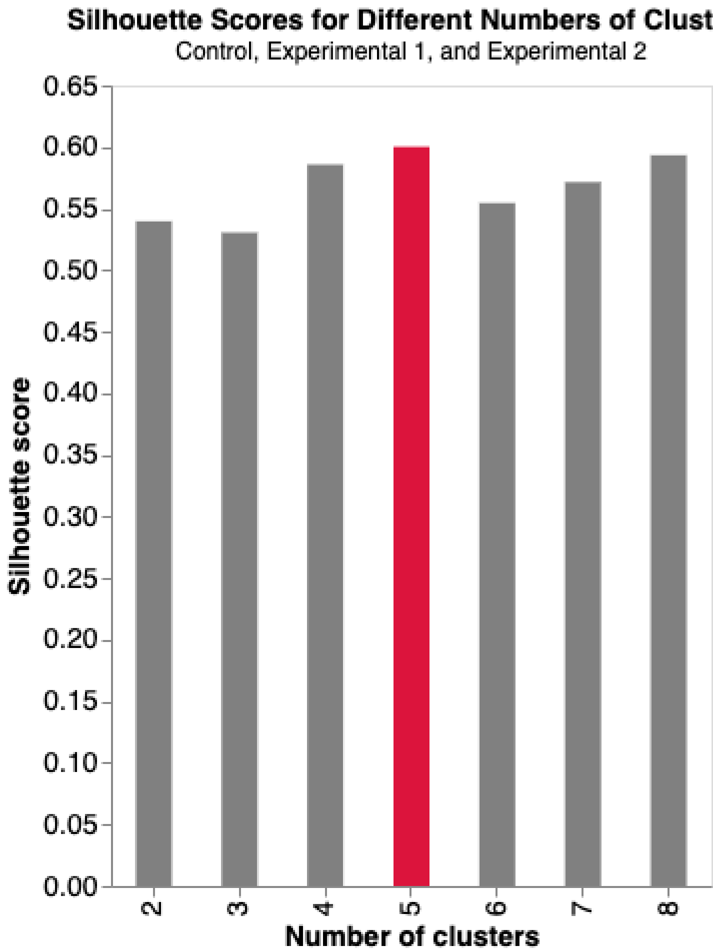


Figure 4.8 Graphical Silhouette Coefficient Clustering for all Groups (N = 60)

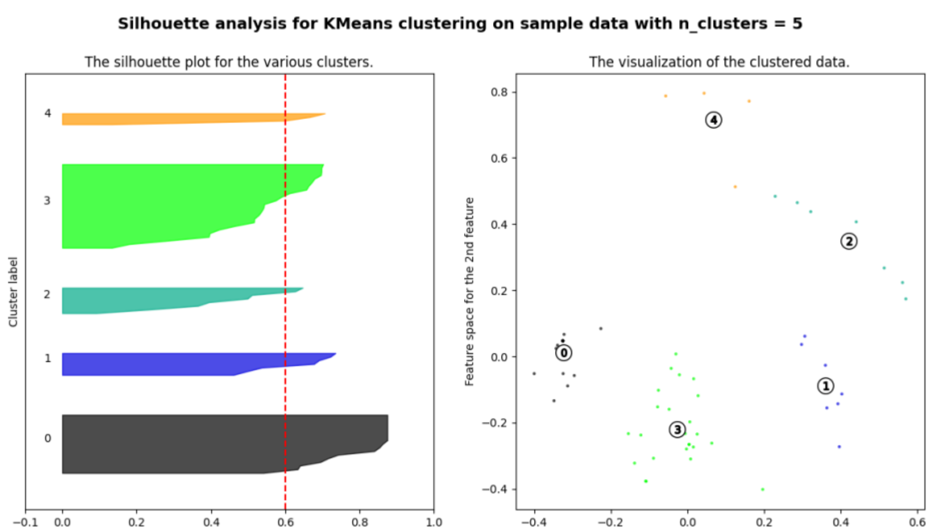


Figure 4.9 K-means Cluster for all Groups (N = 60)

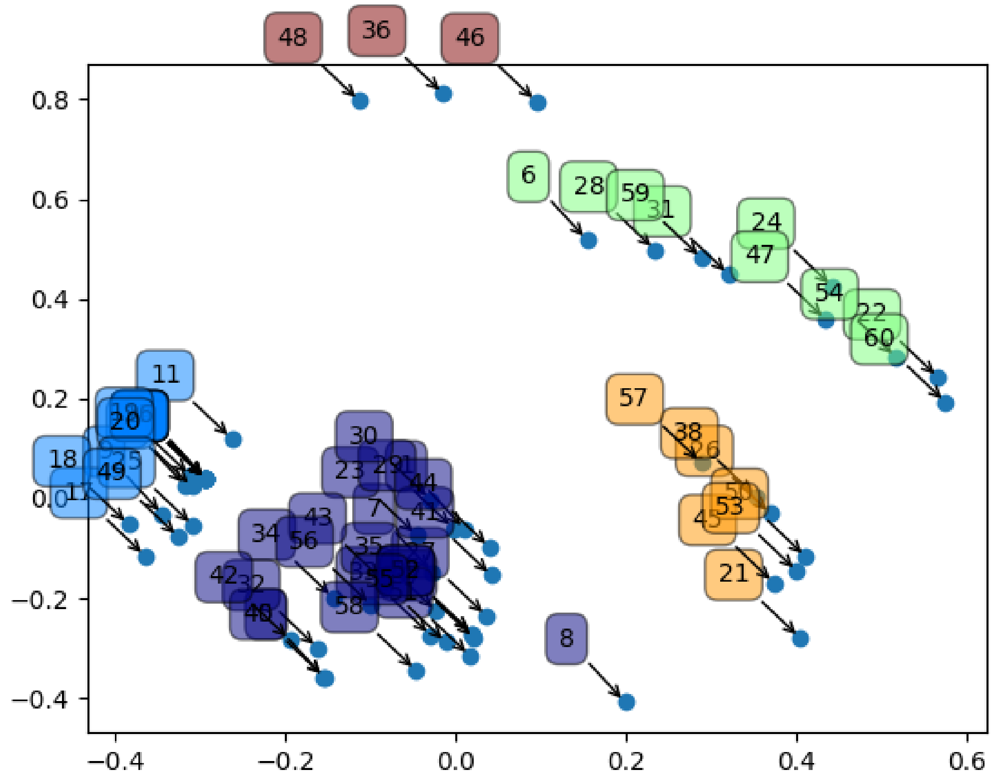


Figure 4.10 shows the median strings (the most representative strings of engineering approaches) for all the clusters from the Control, Experimental 1, and Experimental 2 groups. Figure 4.11 displays the median strings (the most representative strings of search strategies and knowledge integration) for all five clusters identified across the Control, Experimental 1, and Experimental 2 groups.

Figure 4.10 K-means Clusters and Generalized Median Strings for Engineering Approaches of all Groups (N = 60)

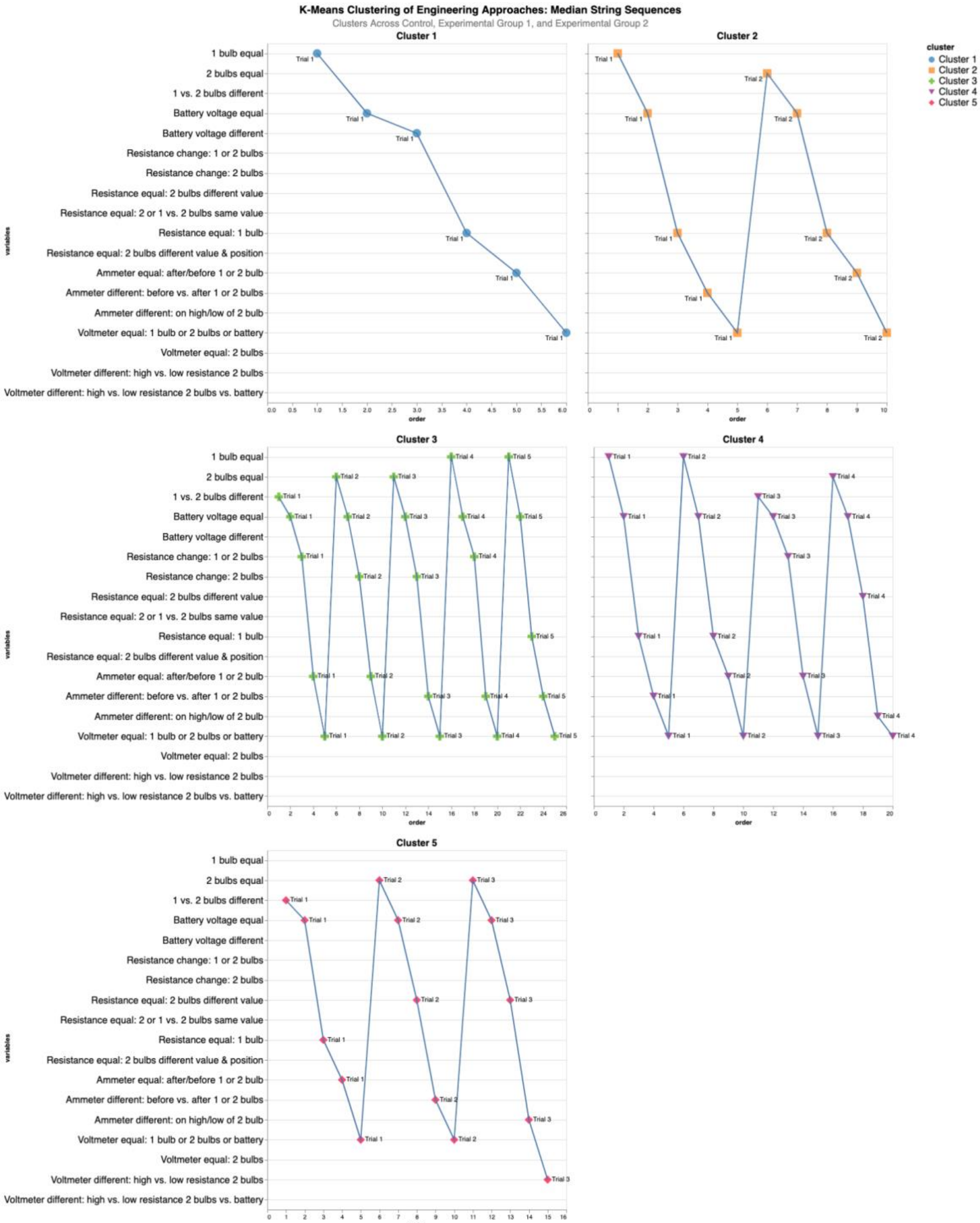
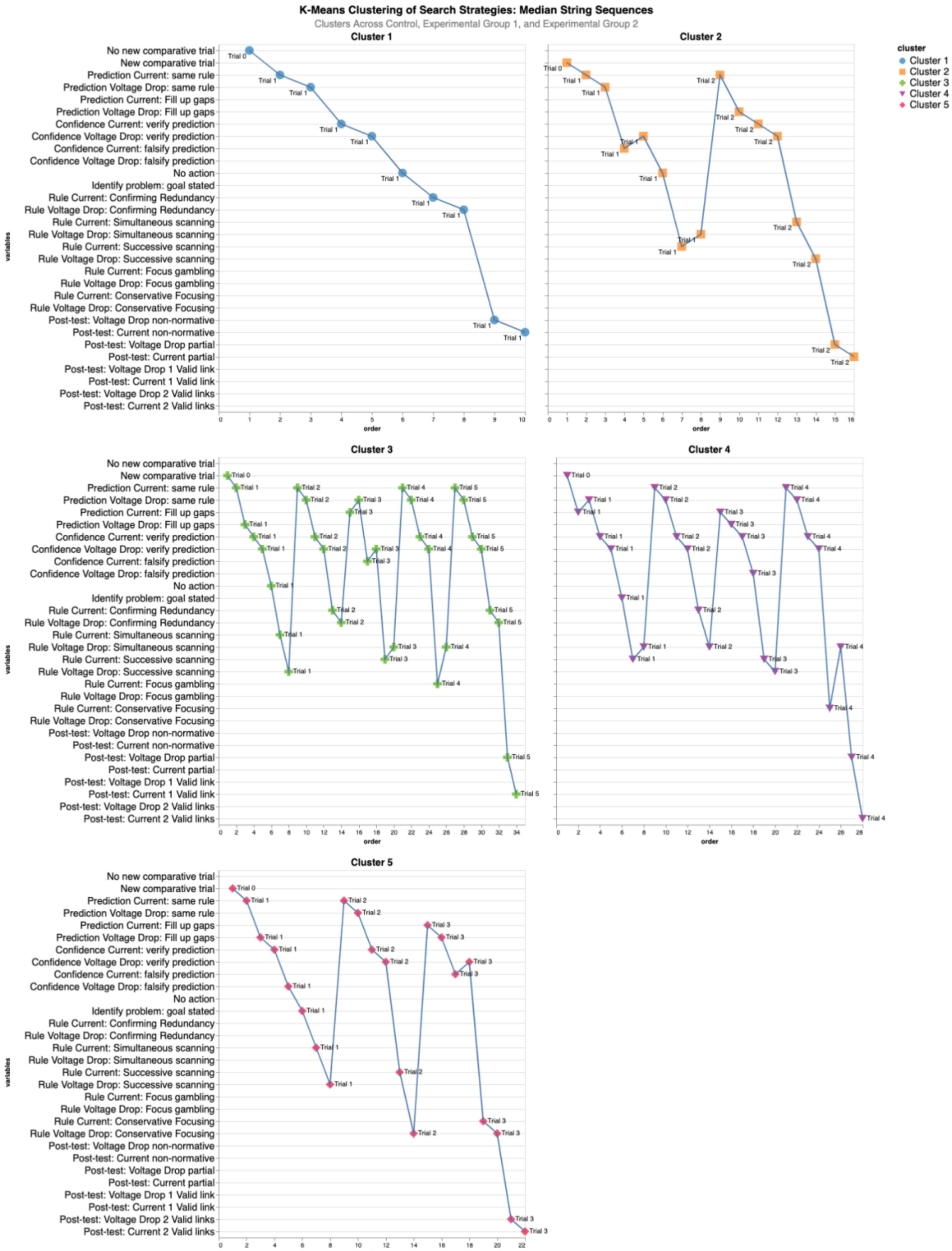


Figure 4.11 K-means Clusters and Generalized Median Strings for Search Strategies of all Groups (N = 60)



4.2.2. Cluster 1

The engineering approaches of cluster 1 are displayed in Figure 4.10, and the search strategies of this cluster are presented in Figure 4.11. In the investigation phase, the participants within Cluster 1 did not initiate any new comparative trials. Instead, they mirrored the trials tested in the exploration phase. Predictions were initially made for the current and voltage drop in the comparative trials, based on rules identical to the ones used in the previous exploration phase trial. As an example, participant A from Cluster 1 expressed, “I will test a circuit with a 10V battery and a 5-Ohm resistor to determine the current and potential difference when I increase the resistance. I predict that the current will decrease, but the potential difference will remain constant.”

Following these predictions, the anticipated current and voltage drop values were verified. It is noteworthy to mention that during this phase of investigation, no specific actions related to problem identification were documented. When forming rules, Cluster 1 consistently used the “confirmation redundancy” strategy for both current and voltage drop. This involved the repeated testing of alterations of the same instance for both parameters in search of hypothesis confirmation.

The post-testing stage revealed non-normative ideas in voltage drop and current, suggesting potential discrepancies that the cluster failed to accurately anticipate. For example, Participant A in Cluster 1 responded to the post-test question by stating, “Current has a negative relationship with resistance; even though my circuit only had one resistor instead of two, increasing the ohms in that resistor decreased the value in the ammeter reading.” Furthermore, they responded to a question comparing Circuit A (with an ammeter before the bulb) and Circuit B (with an ammeter after the bulb), stating, “In my circuit with one resistor, the current dropped as the resistance increased. This means that the reading on the ammeter will be greater before the resistance in the circuit than after the resistance. Given that the current is moving clockwise and the resistors are on the left in Circuit B, this suggests that the current will be less than it was in Circuit A.” Regarding the voltage drop, the participant incorrectly asserted, “The voltage drop is independent of the resistance.” These responses suggested a limited exploration of current and voltage drop in a circuit with one bulb, as opposed to experiments involving multiple bulbs with varying resistance values and different ammeter and voltmeter placements.

4.2.3 Cluster 2

The engineering approaches of cluster 2 are displayed in Figure 4.10, and the search strategies of this cluster are presented in Figure 4.11. During the investigation phase, Cluster 2 participants exhibited a proactive approach by initiating fresh comparative trials. Their initial predictions for the current and voltage drop during the first comparative trial were rooted in principles derived from the preceding stage. For instance, Participant B hypothesized that there would be a change in the current before and after the resistor in a series circuit containing a single resistor. This speculation was based on their earlier explorations, with Participant B noting, "Previous simulation showed that current decreased in proportion with ohm value." These predictions were subsequently subjected to falsification of the current and verification of the voltage drop. Despite their active start, they failed to identify any problems during this stage, a pattern that was also observed in Cluster 1.

In the initial trial, Cluster 2 participants utilized distinct rule generation strategies when examining current. Specifically, they applied a 'successive scanning' strategy, implying an evaluation approach that assessed attributes individually. However, for the voltage drop, they adopted a 'simultaneous scanning' approach, indicating a method that evaluated all attributes collectively and dismissed some based on available evidence. This doesn't imply that the attribute itself (in this case, attributes related to voltage drop) is entirely discarded from the pool of attributes, but rather, it is dismissed as a defining characteristic of the current hypothesis in the initial trial.

In the second trial, the current predictions followed the established rule, while the voltage drop predictions adapted to fill gaps from the previous rule. These new predictions underwent verification, and further rule formation was observed. This time, 'simultaneous scanning' was applied to the current and 'successive scanning' to the voltage drop, an inversion of the strategies used in the first trial.

Post-test results revealed only a partial alignment with the standard norms, indicating some deviation in both current and voltage drop. For instance, Participant B from Cluster 2 established a partially correct link about current, stating, "The placement of the ammeter does not affect the current." However, they inaccurately added, "Adding more resistors to a series circuit leads to a decrease in current and differences in current in the resistors. The resistor with 10 Ohms is less than the resistor with 50 ohms. So the current will be increased in the 10 Ohms conductor as compared to the 50 Ohms conductor." They also made a partially correct statement about voltage drop, indicating, "The voltage drop will be the same as the battery voltage," but incorrectly added, "Resistors do not impact the voltage drop."

In conclusion, Cluster 2 participants employed distinct strategies in rule generation, underscoring varied approaches to managing and predicting current and voltage drop. The observed deviations from expected outcomes in the post-test phase across both clusters suggest potential areas for refining the predictive models or rules.

4.2.4. Cluster 3

The engineering approaches of cluster 3 are displayed in Figure 4.10, and the search strategies of this cluster are presented in Figure 4.11. During the investigation phase, participants in Cluster 3 initiated novel comparative trials distinct from the exploration phase. However, problem identification did not feature in the trials.

In the first trial, they set out to compare the effect of using different numbers of bulbs on resistance. They kept the battery voltage constant and manipulated the placement of the ammeter and voltmeter either before or after the bulbs, or across the battery or bulbs. They also employed previously comprehended rules to predict current and voltage drop, showcasing their ability to apply learned concepts in a new context. These predictions were then verified, affirming their understanding. Throughout these trials, the problem identification did not occur. However, they adopted simultaneous scanning for current and successive scanning for voltage drop, diverse scanning rules that enhanced their understanding of current and voltage drop in the circuit.

In Trial 2, participants modified the setup by placing two identical bulbs in the circuit while keeping the battery voltage constant. The ammeter and voltmeter placement remained the same as in Trial 1. They maintained the same predictive rules as before and verified their predictions, further solidifying their understanding. Here, they introduced confirmation redundancy in their approach, suggesting an investigative method to repeatedly test the same instance.

As they moved into Trial 3, participants adjusted the placement of the ammeter, marking a shift from the previous two trials. While keeping the bulb configuration, battery voltage, and resistance similar to Trial 2, they wanted to understand how ammeter placement affects current readings. They maintained the same predictive rule for voltage drop but transitioned to the “fill up gaps” strategy for current prediction, indicating an effort to address any incomplete understanding from the preceding rule. They confirmed their voltage drop prediction but falsified the current prediction during the confidence measure phase, signifying new insights into the behavior of current within the circuit.

In Trial 4, participants returned to a single bulb configuration, keeping the battery voltage constant while varying resistance. The ammeter was placed differently from the previous trials, continuing the exploration of its impact on current readings. For rule generation, the focus gambling strategy was introduced for current, suggesting concurrent exploration of various attributes related to the current.

Finally, in Trial 5, participants reverted to their initial engineering approaches, predictive rules, and verification strategies. They also returned to employing confirmation redundancy for both current and voltage drop. This pattern implied they were reiterating the rules they found most accurate or consistent throughout their trials. However, their post-test results indicated only a partial understanding of voltage drop and one valid link in comprehending current. For instance, Participant C offered one valid link about current, stating: “No matter where you place the ammeter in the circuit, the reading will always be the same. In two different circuits where the ammeter was placed before and after the bulb, the current before and after the bulbs were the same.” This was juxtaposed with an incorrect understanding of current in a series circuit: “The strength of the current will experience a greater decrease through the first resistor (50 ohms) than through the second resistor because the greater the resistance, the greater the decrease in strength of current flowing out of the end of the resistor. I did not get the chance to test this through the trials.” For the voltage drop, Participant C from Cluster 3 offered a partial link: “The voltage drop is unaffected by the factor of resistors as the value is the same as the battery, and the battery is unaffected", which is partially accurate because they realize the total resistance of two bulbs with different resistance value will be equal to the voltage of the battery. However, they wrongly assumed that the voltage drop for each bulb with different resistance value would be identical, adding: “I'm not quite sure, given that I didn't test what would happen to the voltage drop if you only increased the resistance.” This underscores the students' capability to establish correlations between varied concepts, albeit full integration of knowledge necessitates further cultivation.

Both clusters 3 and 2 showed an understanding of the rules of predicting and verifying the current and voltage drop in the electric circuit simulation. However, Cluster 3 demonstrated a more diverse application of rule generation strategies and a higher level of knowledge integration in the post-test, indicated by the presence of a valid link. On the contrary, Cluster 2 demonstrated an approach that involved addressing gaps in predictions, indicating a more iterative learning process, but their post-test results indicated only partial knowledge integration. Future instructional strategies should aim to foster knowledge integration and encourage diverse rule generation strategies in students to enhance their understanding of electric circuits.

4.2.5. Cluster 4

The engineering approaches of cluster 4 are displayed in Figure 4.10, and the search strategies of this cluster are presented in Figure 4.11. During the investigation phase, Cluster 4 participants engaged in a series of comparative trials, demonstrating distinctive strategic modifications from the exploration phase.

During Trial 1, the participants adeptly employed existing knowledge to predict current, employing the "fill up gaps from preceding rule" strategy. For example, Participant D resolved to examine the current both before and after the resistor, as opposed to placing the ammeter strictly before or after the resistor as in the exploration phase. Participant D predicted that the current would remain consistent before and after a resistor in a series circuit with a single bulb due to constant resistance. Concurrently, they adopted the preceding rule for predicting the voltage drop. The subsequent verification of both the current and voltage drop predictions reinforced confidence in their hypotheses. In terms of rule generation, the participants employed successive scanning for the current and simultaneous scanning for the voltage drop, indicating an efficient process of hypothesis elimination.

In Trial 2, participants maintained the same predictive and verification approaches for current and voltage drop as in Trial 1. However, a strategic shift in the rule generation phase was observed as they transitioned to confirmation redundancy for current and simultaneous scanning for voltage drop. The use of confirmation redundancy was to affirm their understanding of the impact of ammeter placement on the current within a series circuit with a single bulb. This strategic interchange suggests the participants' experimentation with different approaches to decipher patterns or relationships in the data.

In Trial 3, a divergence from the previous two trials was observed as participants engaged in “filling up gaps” and falsified the current and voltage drop predictions, indicating an openness to reevaluating assumptions. During this trial, the number of bulbs was increased to compare a single bulb with two bulbs in a series circuit, keeping the ammeter placement constant while altering the voltmeter placement. The participants predicted that the current in a circuit with two bulbs would vary due to increased overall resistance. In tandem, they predicted no change in the voltage drop when the voltmeter is placed across one bulb versus two bulbs. However, after testing their hypothesis, they falsified their prediction. The use of successive scanning was evident for both current and voltage drop during rule generation, demonstrating a continued engagement with hypothesis testing and falsification.

By Trial 4, the participants refined their approach further, testing the current on higher and lower resistance bulbs within a series circuit encompassing two bulbs with different resistance values. Predictions and verifications proceeded as before, while the rule generation phase was marked by the adoption of conservative focusing for current and simultaneous scanning for voltage drop.

Post-test evaluations showcased participants’ partial comprehension of voltage drop, while they established two valid links for current. For instance, Participant D acknowledged, “In a series connection, the current remains the same irrespective of the positioning of the ammeter,” and they remarked that "the current across the resistor stays the same even if the value of resistance changes in a series connection. “Nonetheless, Participant D demonstrated a partial understanding of voltage drop, correctly stating initially that the “voltage drop was equal to the amount of volts in the battery,” but later providing incorrect assertions about the voltage drops across resistors of different resistances.

Participants in Cluster 4 exhibited flexibility and self-regulation across the trials, grappling with the complexities of understanding current and voltage in electric circuits. This was reflected in their evolving search strategies and engineering approaches. However, incomplete links of ideas, particularly relating to voltage drop across resistors of varying resistances, point to areas where further instruction or practice could be beneficial.

4.2.6. Cluster 5

The engineering approaches of cluster 5 are displayed in Figure 4.10, and the search strategies of this cluster are presented in Figure 4.11. In Cluster 5, the students conducted new comparative trials. In the first trial of the investigation phase, the students prediction regarding current was derived from preceding rules, indicating an ability to build upon previous knowledge. However, their prediction regarding voltage drop changed from the preceding rule in the exploration phase. For instance, Student E predicted that: “there is a difference in voltage drop” when manipulating the number of bulbs from one to two and placing the voltmeter across one bulb versus two bulbs. After testing this prediction, they found there was no change in the voltage drop. Confidence measures post-simulation illustrated their ability to both affirm and challenge their predictions. They verified their prediction for the current, while concurrently demonstrating critical reasoning by falsifying their prediction for the voltage drop. The students’ capacity to revise and align their understanding based on the evidence at hand was thus demonstrated. Additionally, they identified a specific problem, signifying their capacity to focus their learning around particular objectives or challenges. In terms of rule generation strategies, the students adopted a differentiated approach for the current and voltage drop. The strategy of simultaneous scanning was employed for the current, wherein all attributes were considered concurrently, facilitating the elimination of certain hypotheses. For the voltage drop, they adopted the successive scanning method, focusing on one attribute at a time.

In the subsequent trial, the students consistently affirmed their predictions for both the current and voltage drop, indicating their progressive understanding of the electric circuit's behavior. The rule-generation strategies for both current and voltage drop showed a change in one attribute at a time (i.e., successive scanning), implying a degree of consistency in their approach.

In the final trial, the students continued their prediction strategy to fill the gaps from the preceding rule. For instance, Participant E tried to bridge the gap in the previous rule by predicting “there is a difference in voltage drop because I am changing the value of each bulb” and predicting “there is a difference in current because I am changing the resistance of each bulb” in a circuit with 2 bulbs. Unlike the other cluster of students, they identified problems as to why they did not observe a change in voltage drop when they manipulated the resistance of 1 bulb in a series circuit; how they may change the voltage drop in a series circuit; and how they may change the current in different locations and resistances of a series circuit. For example, Participant E identified the goal, sub-goals, and obstacle to achieve the goal state:

> Wait, how would I plot this if I wanted to change the resistance? [goal] Okay so if two bulbs, each set at a different value, current after both of the bulbs will be decreased because it’ll go through the first one and then decrease [sub-goals]. Oh well, it’s gonna decrease no matter what but I don’t know if it’ll decrease more than it already does I guess [obstacle].

Concurrently, they demonstrated a measure of discernment in their confidence by falsifying their prediction for the current and verifying the voltage drop. A notable evolution in their rule-generation strategy was observed, shifting towards conservative focusing. This change involves investigating a hypothesis by altering a single attribute on each trial, thus reflecting a more systematic approach to problem-solving.

The students' post-test responses indicated their successful knowledge integration. They exhibited the ability to make multiple valid links concerning the current and voltage drop. For instance, Participant E made multiple valid links about the current, saying, "Current does not change from its relative position between the resistor and the battery. Current is the same throughout a circuit regardless of its position according to the number of resistors present and the value of voltage in the formula Current = Voltage/Resistance. The current will drop for the whole circuit, but remains the same for each resistor in the circuit." Additionally, Participant D provided multiple valid links about the voltage drop, saying, "The voltage drop would be the same among the 2 light bulbs assuming the resistance and the voltage remain constant but if the light bulbs had different resistance then the voltage drops would be different. The voltage drop across all sources of resistance will equal 30 volts because that is what the battery provides. Voltage drop is affected by resistance, if there is more resistance the voltage drop will be higher. Voltage drop across a resistor is equal to the ratio of that resistors' resistance to the total resistance in the series circuit (40/60 > 20/60)."

In sum, students in Cluster 5 demonstrated a progression in their problem-solving abilities, displaying consistency and flexibility in adjusting their strategies based on their findings.

4.2.7. Sequential Analysis within Each Group (Control, Experimental 1, and Experimental 2)

I analyzed the patterns in the sequential use of search strategies among different clusters within Experimental group 1, Experimental group 2, and the Control group. The sequential analysis of the three clusters presented intriguing findings on the use of strategies, such as prediction, confidence measures, no action vs. problem identification, and rule formation.

Figure 33 shows the Silhouette scores against the values of clusters for the Control group. Among different clusters, k = 3 has the highest silhouette score average while k = 2 has the lowest silhouette score average (Figure 4.12). Among different clusters k = 3 has the least negative values. Figure 4.13 is the graphical Silhouette Coefficient clustering that shows that k = 3 has no negative values. Thus, I selected k = 3 as the optimal number of clusters. After testing different numbers of clusters, 3 clusters showed more inter-cluster similarity and intra-cluster dissimilarity (Figure 4.14).

Figure 4.12 Silhouette Scores for Control Group (N = 20)

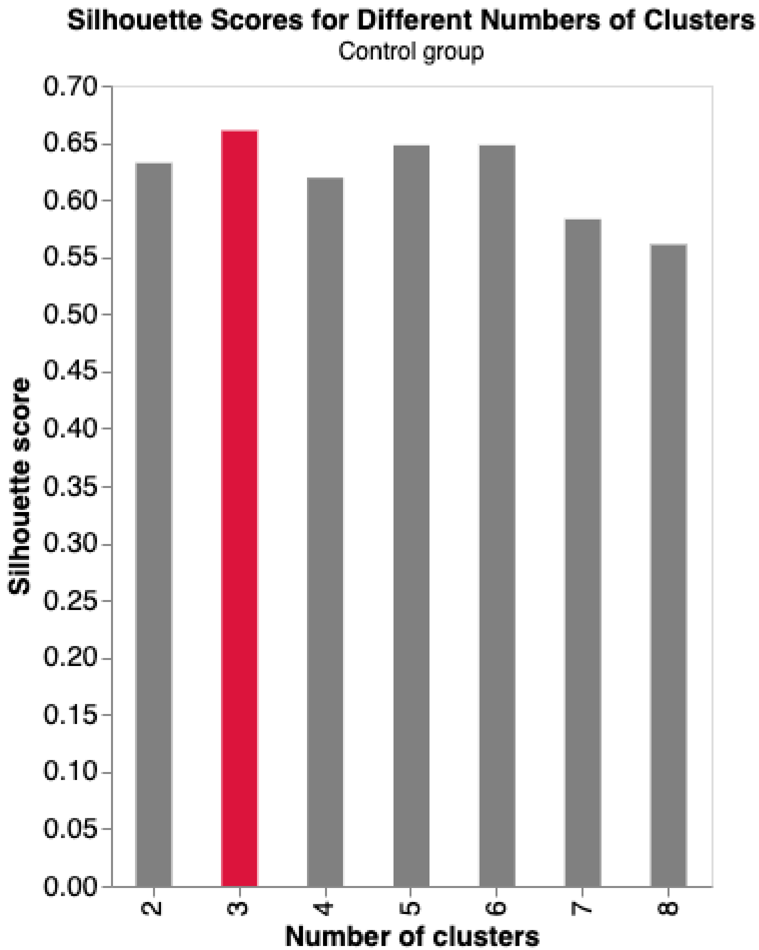


Figure 4.13 Graphical Silhouette Coefficient Clustering for Control Group (N = 20)

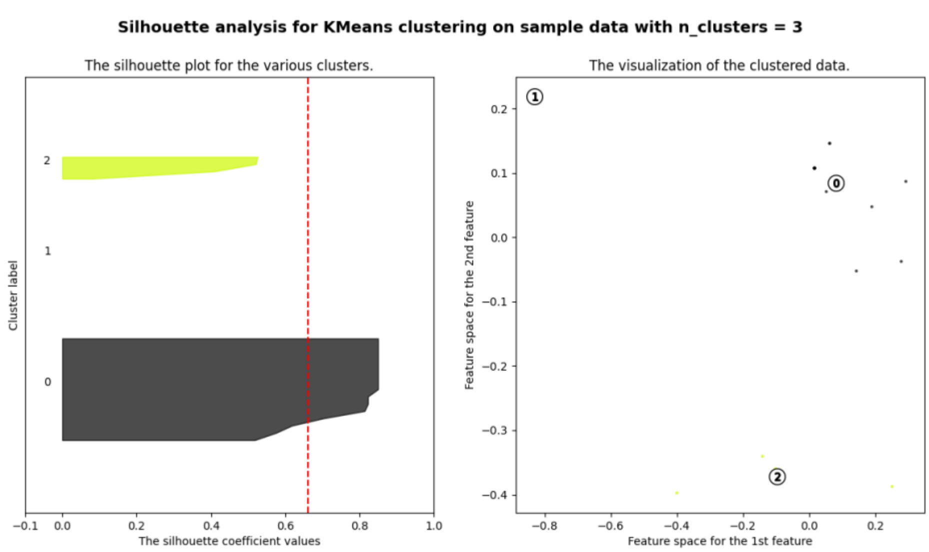


Figure 4.14 K-Means Cluster for Control Group (N = 20)

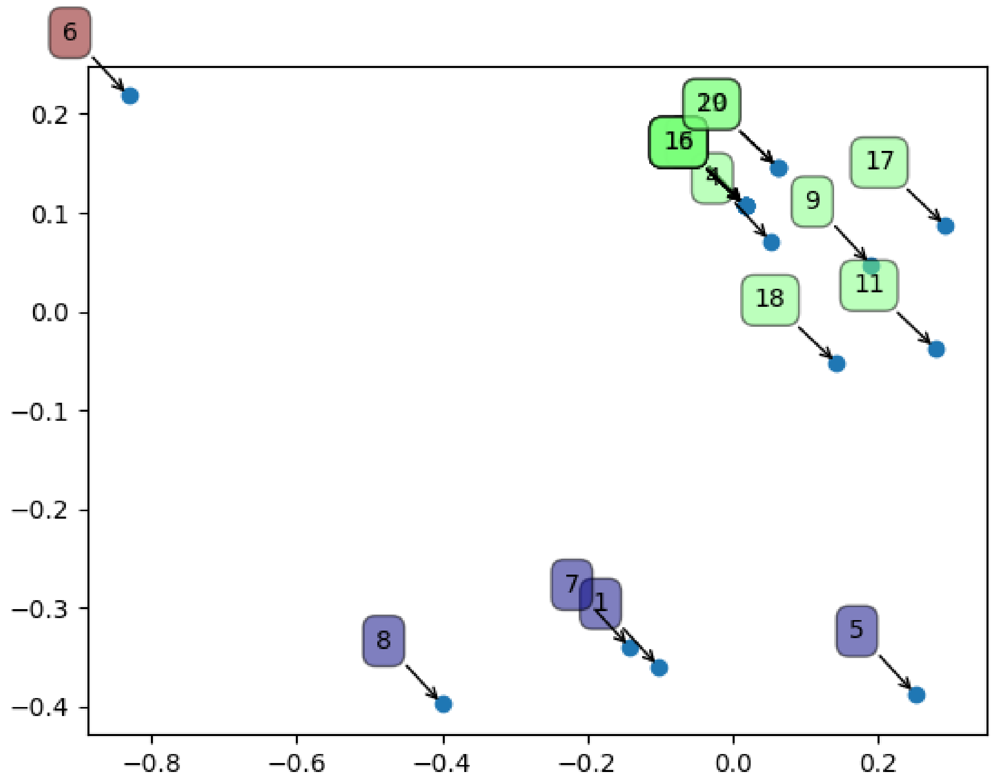


Figure 4.15 shows the Silhouette scores against the values of clusters for Experimental 1 group. Among different clusters, k = 3 has the highest silhouette score average while k = 2 has the lowest silhouette score average (Figure 4.15). Among different clusters k = 3 has the least negative values. Figure 4.16 is the graphical Silhouette Coefficient clustering that shows that k = 3 has no negative values. Thus, I selected k = 3 as the optimal number of clusters. After testing different numbers of clusters, 3 clusters showed more inter-cluster similarity and intra-cluster dissimilarity (Figure 4.17).

Figure 4.15 Silhouette Scores for Experimental 1 Group (N = 20)

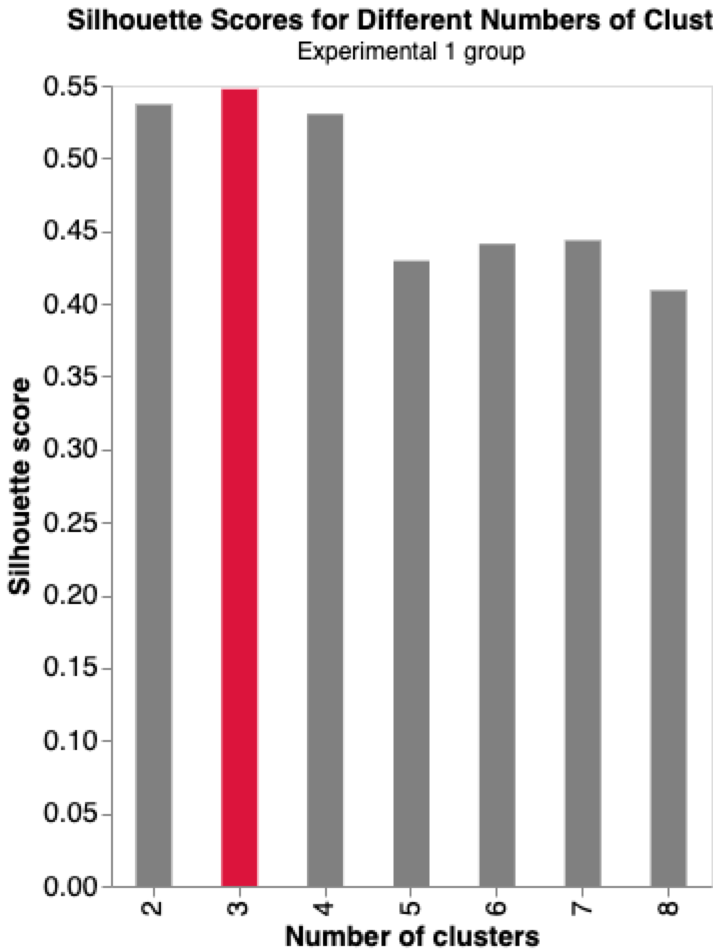


Figure 4.16 Graphical Silhouette Coefficient Clustering for Experimental 1 Group (N = 20)

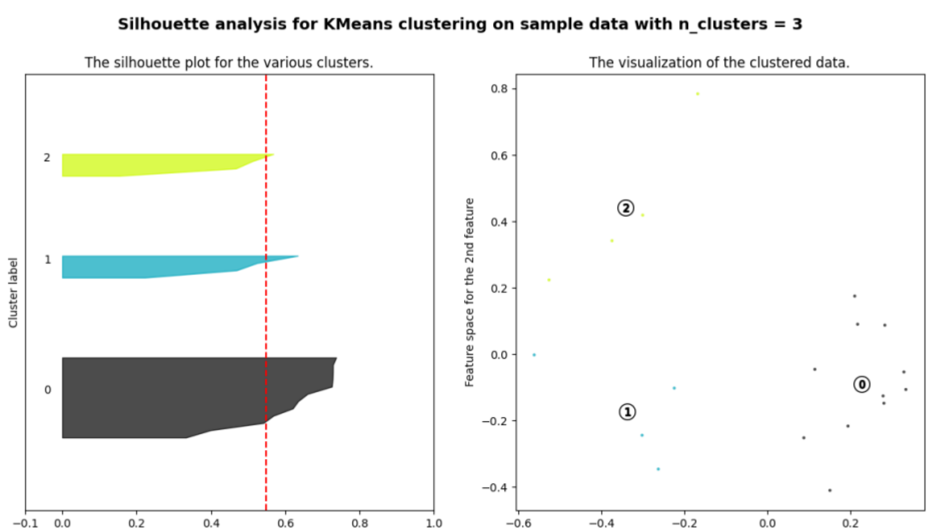


Figure 4.17 K-Means Cluster for Experimental 1 Group (N = 20)

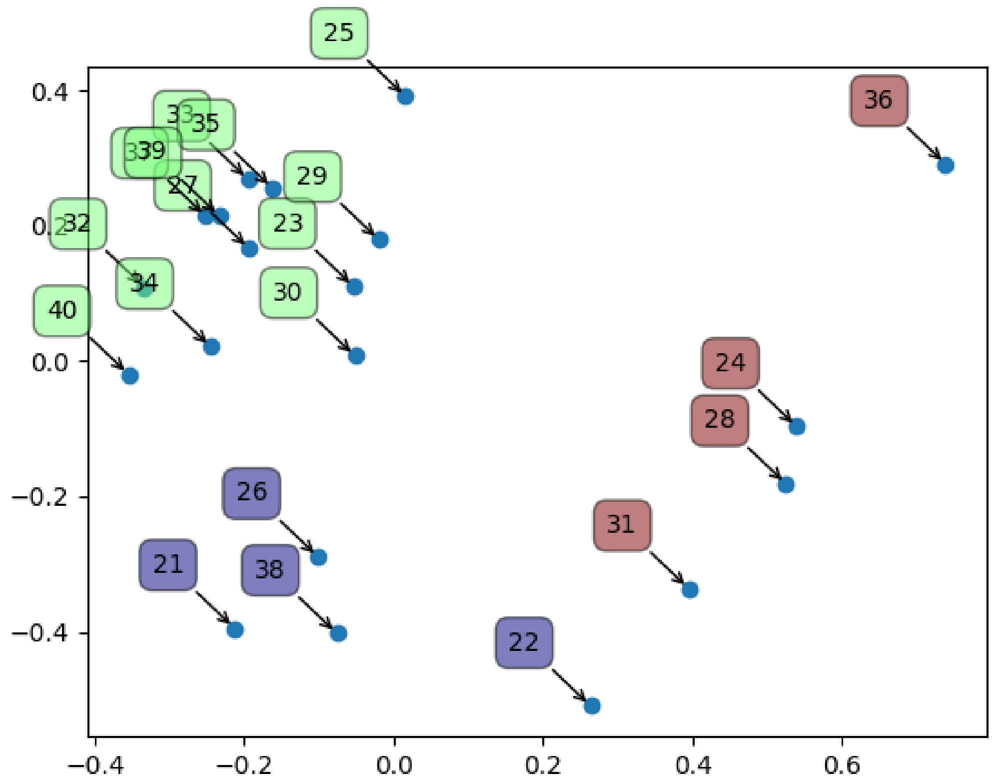


Figure 4.18 shows the Silhouette scores against the values of clusters for Experimental 1 group. Among different clusters, k = 3 has the highest silhouette score average while k = 2 has the lowest silhouette score average (Figure 4.19). Among different clusters k = 3 has the least negative values. Figure 40 is the graphical Silhouette Coefficient clustering that shows that k = 3 has no negative values. Thus, I selected k = 3 as the optimal number of clusters. After testing different numbers of clusters, 3 clusters showed more inter-cluster similarity and intra-cluster dissimilarity (Figure 4.20).

Figure 4.18 Silhouette Scores for Experimental 2 Group (N = 20)

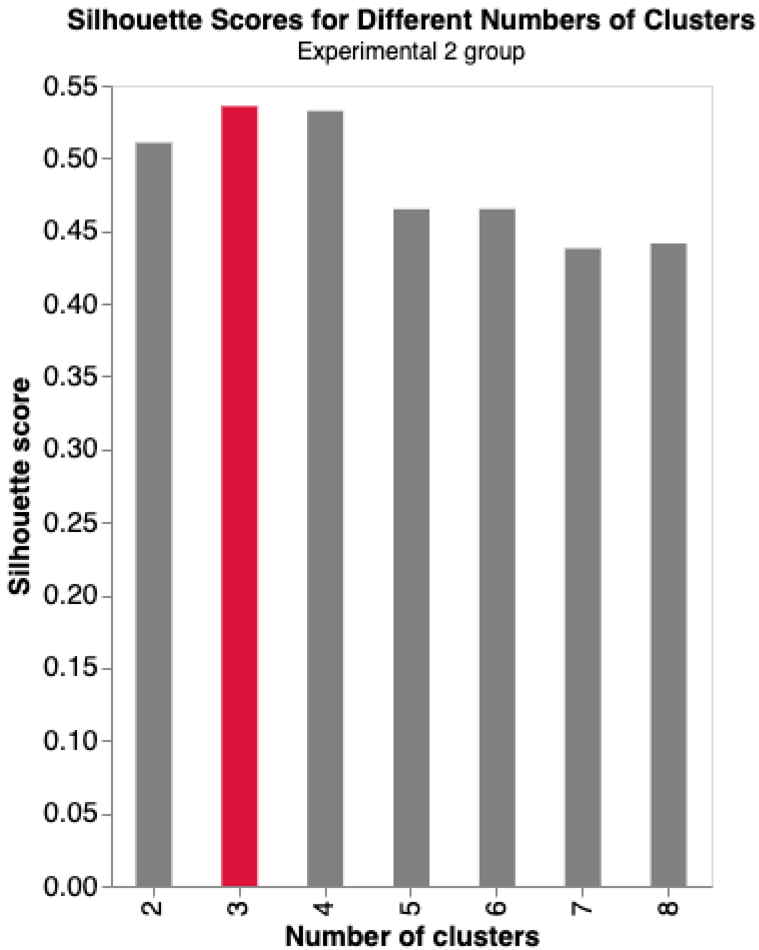


Figure 4.19 Graphical Silhouette Coefficient Clustering for Experimental 2 Group (N = 20)

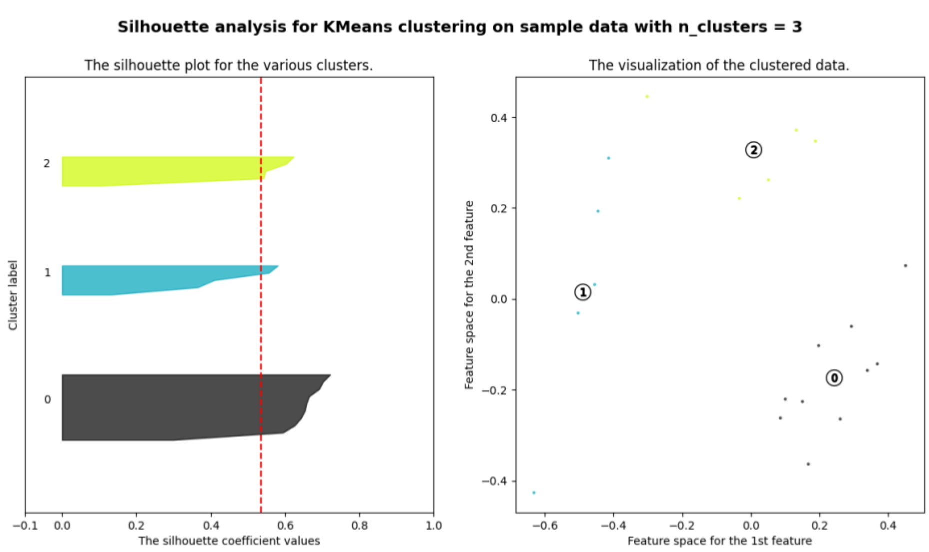
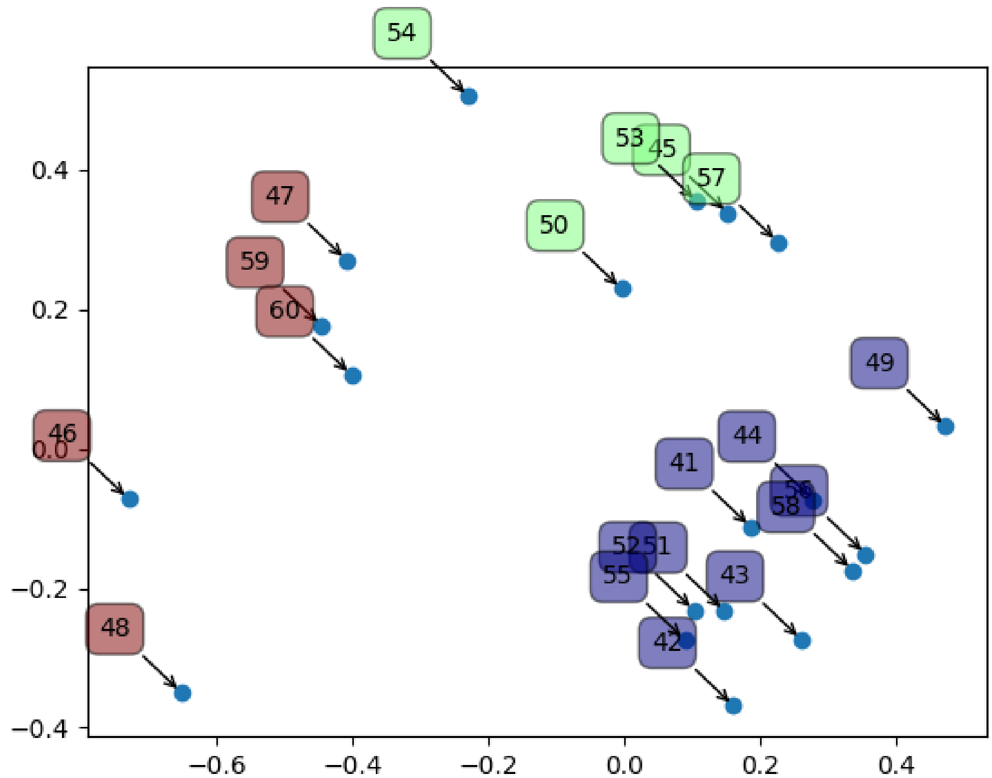


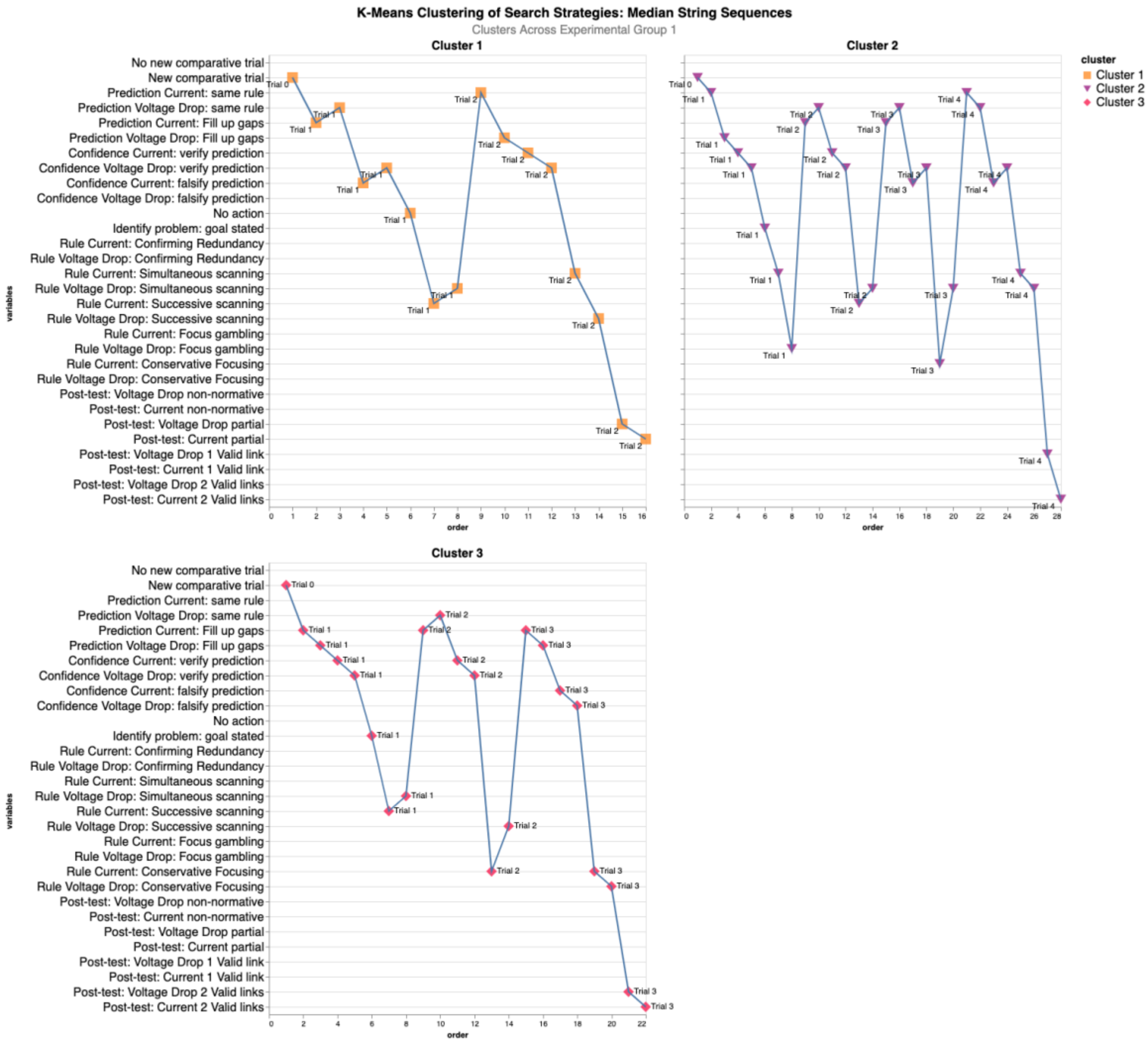
Figure 4.20 K-Means Cluster for Experimental 2 Group (N = 20)



4.2.7.1. Experimental Group 1

Figure 4.21 displays the clusters in Experimental group 1. In this group, Cluster 1 displayed a tendency towards 'Fill up gaps' in the first trial, followed by 'same rule' in the subsequent trial. 'Falsify' and 'verify' were interchangeably used as confidence measures. 'Successive scanning' and 'Simultaneous scanning' appeared as the main rule formations. Similar patterns were observed in Cluster 2 and 3 with noticeable differences in problem identification and rule formation strategies. For instance, Cluster 2 incorporated 'Identify problem: goal stated' in trial 1 and used a variety of rule formation strategies such as 'Focus gambling' and 'Conservative Focusing'. In Cluster 3, a consistent 'Fill up gaps' prediction approach was utilized across all trials with an alternating confidence measure strategy. In the post-test, Cluster 1 showed a partial understanding of Voltage Drop and Current. Cluster 2 managed to establish one valid connection for Voltage Drop and two for Current. Meanwhile, Cluster 3 was able to construct two valid connections each for Voltage Drop and Current.

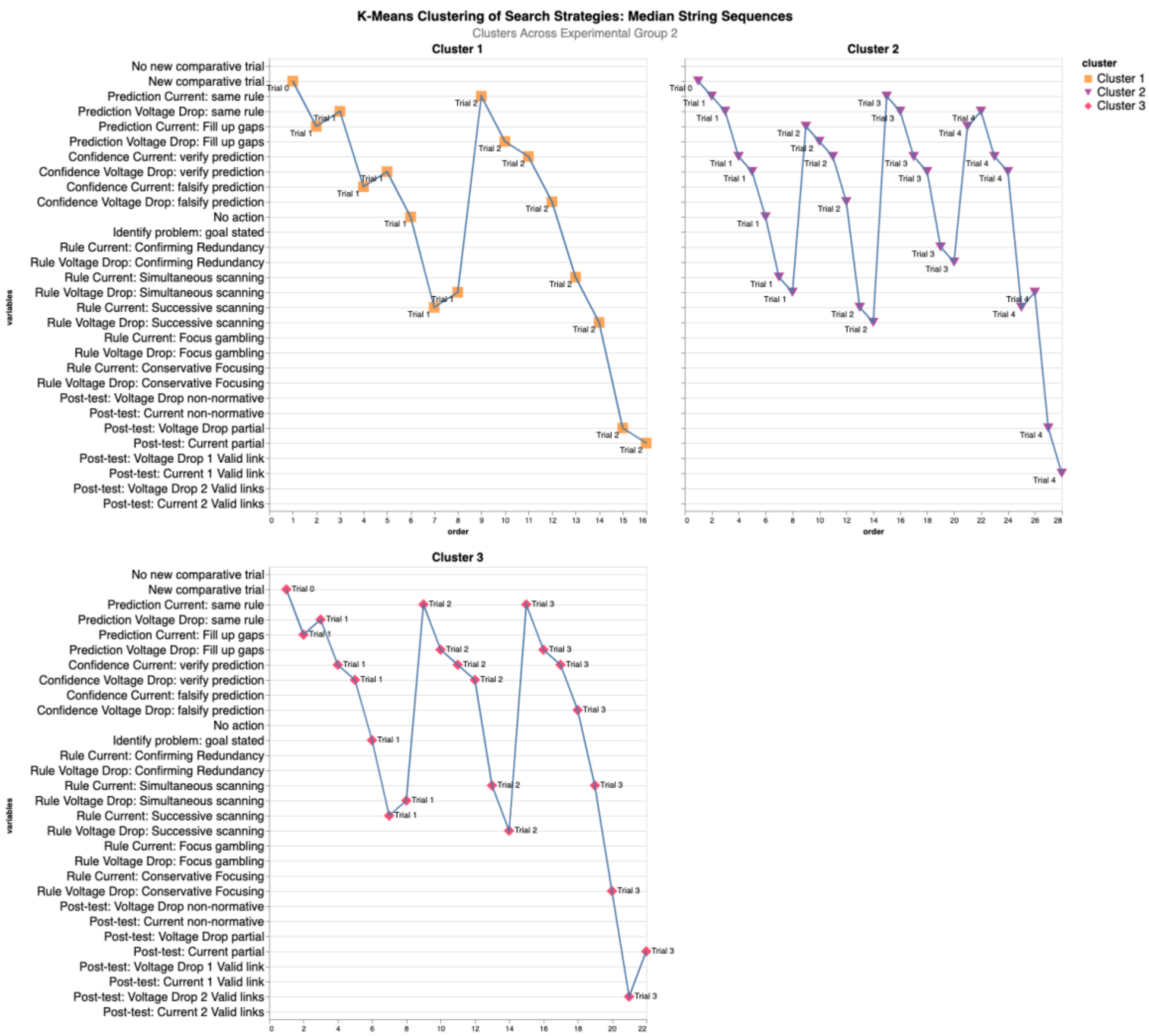
Figure 4.21 K-Means Clusters and Generalized Median Strings for Search Strategies of Experimental 1 (N = 20)



4.2.7.2. Experimental Group 2

Figure 4.22 showcases the clusters in Experimental group 2. The patterns observed in Experimental group 2 deviated slightly from those in group 1. Cluster 1 from group 2 paralleled Cluster 1 from group 1, although the confidence measures shifted in the second trial where 'falsify' was specifically applied to 'Voltage Drop'. Cluster 2 showcased a more consistent prediction strategy with 'same rule' and 'Fill up gaps' being applied interchangeably across trials. In addition, 'Confirming Redundancy' was employed as a rule formation strategy in trial 3. Cluster 3, interestingly, witnessed a similar approach as Cluster 1 with an extra use of 'Conservative Focusing' in the final trial. During the post-test, Cluster 1 only exhibited a partial understanding of Voltage Drop and Current. Cluster 2 also demonstrated a partial grasp of Voltage Drop but managed to form one valid connection for Current. In contrast, Cluster 3 succeeded in forming two valid links for Voltage Drop, but showed only a partial comprehension of Current.

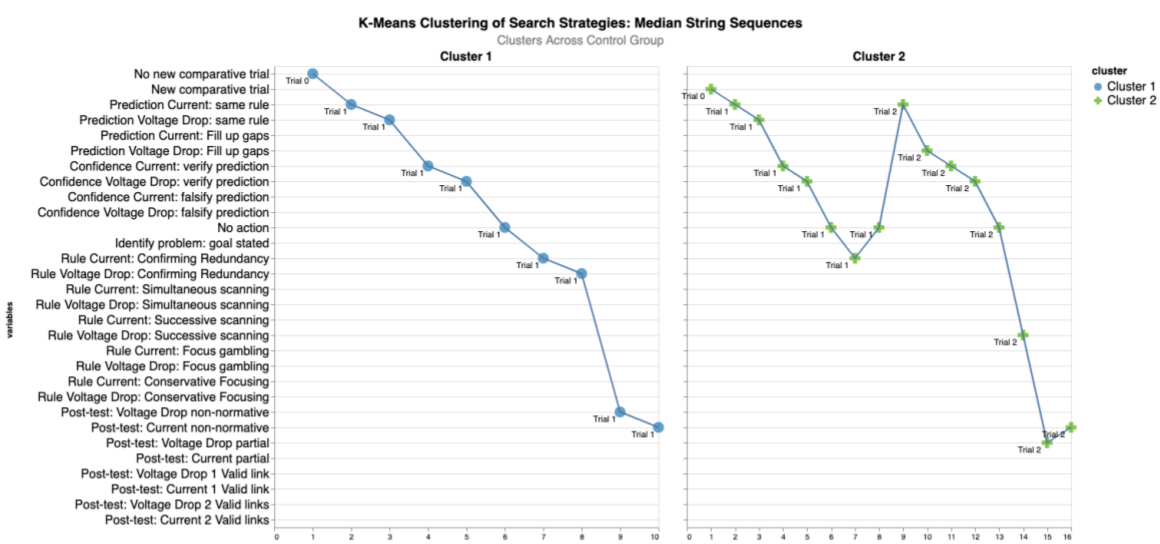
Figure 4.22 K-Means Clusters and Generalized Median Strings for Search Strategies of Experimental 2 (N = 20)



4.2.7.3. Control Group

Figure 4.23 presents the composition of clusters within the Control group. The Control group, however, exhibited a more pronounced difference in their search strategy patterns. Cluster 1 did not undertake a new comparative trial in trial 0, which differed from Experimental groups. The 'same rule' approach was adopted in prediction while 'verify' was consistently used for confidence measures. Moreover, 'Confirming Redundancy' was the predominant rule formation strategy. This differs significantly from Cluster 2 where 'No action' was repeatedly taken, suggesting a lack of active search strategies employed by this cluster. The empirical analysis suggests varied patterns of strategy usage across the clusters within each group. Some clusters favored certain strategies, such as 'same rule' for prediction, while others exhibited more variability. The presence of unique approaches within clusters, such as the 'No action' strategy observed in the Control group, further emphasizes the diversity of search strategies within these groups. These findings shed light on the dynamic interplay of strategies and provide insights into how these methods are applied and alternated in different experimental settings. Future research should consider these patterns to understand the factors influencing these strategic choices and to further explore the impact of these strategies on the outcomes of the tasks undertaken. In the post-test, Cluster 1's understanding of Voltage Drop and Current was classified as non-normative. Cluster 2 exhibited a partial understanding of Voltage Drop, while its comprehension of Current was also classified as non-normative.

Figure 4.23 K-Means Clusters and Generalized Median Strings for Search Strategies of Control Group (N = 20)



Chapter 5. Discussion

The primary objective of this study was to compare across various variables the performance of two experimental groups, namely Experimental 1 using both the decision table and rule induction tools and Experimental 2 using only the rule induction tool, with a Control group. The study's focus was investigating the progression of engineering approaches, search strategies, and the integration of ideas among students using an electric circuit simulation under different conditions: Control, decision table and induction rule tool, and induction rule tool conditions. The results obtained from statistical analyses, incorporating Dunn's test and Spearman correlation coefficient, offered insights into the impacts of distinct tools and techniques on engineering approaches and search strategies. Furthermore, employing sequence analyses, specifically K-means clustering based on Levenshtein edit distance and Levenshtein median, unveiled distinct learning paths as students interacted with electric circuit simulations and different scaffolding tools. Consequently, the findings elucidated significant disparities and correlations within various variables across these conditions, shedding light on students' learning processes and strategies as they engaged with the simulation.

It is imperative to note that the assumption of data normality required for ANOVA was not met, as indicated by the Shapiro-Wilk test for normality of residuals. Consequently, this led to the utilization of non-parametric tests, specifically Dunn's test, as an appropriate alternative. This decision was made due to the violation of normality assumptions and the presence of ordinal data. Dunn's test, in this context, facilitated accurate mean rank comparisons among the three groups.

5.1. RQ 1. What progress in engineering approaches, search strategies, and integrating ideas do students make while using the electric circuit simulation in the control condition versus rule formation conditions?

Research Question 1 (RQ1) concerns students' uses of engineering approaches, search strategies, and the integration of ideas when using electric circuit simulations under different instructional conditions. Results revealed noteworthy patterns and statistically detectable differences between the Control group and Experimental groups 1 and 2 across various dimensions, including prediction, confidence measures, problem identification, rule formation, and post-test performance. These findings can be meaningfully contextualized within the framework of existing educational theories and models.

In terms of problem identification and action variables, marked differences were observed between the Control group and Experimental group 1 using both a decision table and an induction rule tool. Under the 'No Action' condition, which signifies a lack of engagement in activities like problem identification, prediction, and confidence measurement, there were significant disparities between these groups. This variance suggests a considerable gap in the initiation of new comparative trials during the investigation phase, highlighting a divergence in active learning behaviors between the Control and Experimental group. These insights are consistent with Bruner's theory of discovery learning, which argues that active involvement in learning is crucial for knowledge acquisition (Bruner, 1961). These ideas gain further support from Winne's Self-Regulated Learning (SRL) model (2023), which emphasizes the vital role predictive tools play in facilitating students' learning process (Winne, 2022).

The analysis also unveiled intriguing contrasts when comparing Experimental group 1, equipped with both a decision table and an induction rule tool, to Experimental group 2, which had only the induction rule tool at its disposal. In the context of action variables, both Experimental groups 1 and 2 showed significant differences in the "No Action" variable, thereby indicating varying levels of passive behavior between these groups.

Overall, these patterns contribute to understanding how different instructional tools and conditions affect students' engagement and learning outcomes, providing valuable implications for educational practices. Examining prediction variables revealed important nuances in learning behavior across groups. Specifically, in the context of voltage drop, both 'Same rule' and 'Fill up gaps' variables exhibited detectable differences between the Control and Experimental group 1, highlighting divergent approaches to problem-solving. However, for electrical current, only the 'Fill up gaps' variable revealed a discernible difference, pointing to a detectable distinction in prediction strategies. When focusing on Experimental group 2, which had access to only the rule induction tool, distinct patterns in search strategies emerged. For instance, when predicting electrical current, both 'Same rule' and 'Fill up gaps' showed marked disparities between this group and the Control group. Yet, for voltage drop, the difference was confined to the 'Fill up gaps' variable. A comparative analysis between Experimental groups 1 and 2 further illustrated unique patterns: 'Prediction Current: Same Rule' and 'Prediction Current: Fill up Gaps' were significantly distinct, suggesting varied strategic orientations or understandings concerning current prediction. These results indicate that Experimental group 1, equipped with both a decision table and a rule induction tool, demonstrated more active and precise prediction strategies. This is congruent with Winne's 2023 Self-Regulated Learning (SRL) model, thereby substantiating the idea that certain instructional tools can significantly enhance the prediction phase of student learning (Winne, 2022).

Upon assessing confidence measures, both 'Verifying' and 'Falsifying' conditions manifested detectable differences between the Control group and both Experimental groups 1 and 2. Specifically, for both current and voltage drop, these conditions exhibited significant disparities between the Control and Experimental groups. Nonetheless, the two experimental groups demonstrated no statistically detectable variations in confidence measures between them. These observations resonate with Winne's SRL model, asserting that cognitive tools can empower students to better regulate their cognitive processes, which in turn impacts their confidence (Winne, 2022). Such findings align with a principle from cognitive psychology which suggests that cognitive tools and strategies can significantly influence learners' cognitive processes and their confidence in those processes (Kim & Reeves, 2007). Therefore, even a single tool, such as the rule induction tool, can have a considerable effect on learners' confidence. Consequently, regardless of the distinct tools available to them, both experimental groups manifested comparable confidence levels, reinforcing the insights from Winne's SRL model.

In the domain of rule formation, both 'Confirming Redundancy' and 'Conservative Focusing' variables presented discernible differences between the Control and Experimental group 1, particularly in the context of current. In relation to voltage drop, only the 'Confirming Redundancy' condition exhibited a significant difference. These observations align with Winne's COPES model, which emphasizes the role of "If-Then" outcome expectations and deliberate practice in learning (Winne, 2022; Ericsson &

Harwell, 2019). When comparing rule formation strategies between the Control and Experimental group 2, distinct patterns emerged, particularly in the strategies of 'Confirming Redundancy' and 'Simultaneous Scanning' for both current and voltage drop. Additionally, the comparison between Experimental groups 1 and 2 demonstrated variations in several rule formation variables. These differences were particularly pronounced in variables such as "Identify Problem: Goal Stated," "Rule Current: Simultaneous Scanning," "Rule Voltage Drop: Simultaneous Scanning," and "Rule Current: Conservative Focusing." These divergences indicate that the two experimental groups employed distinct strategies in these specific aspects of rule formation, underscoring the role that tools play in facilitating deliberate practice crucial for the development of expertise (Ericsson & Harwell, 2019). Importantly, these findings lend support to Bruner's argument that the use of more sophisticated tools, such as the decision table and rule induction tool, can enrich learning by fostering a more nuanced understanding of rule formation strategies (Bruner, 1961).

Transitioning to the post-test phase, the study revealed compelling differences in multiple variables between the Control and Experimental groups. Specifically, for the variables 'Non-normative' and 'Two valid links' in both current and voltage drop, notable distinctions emerged between the Control and Experimental group 1. When comparing the Control and Experimental group 2, differences were apparent in the 'Non-normative' and 'Partial' conditions, as well as with '1 Valid Link' for current. However, the '2 Valid Links' condition showed no significant difference. For voltage drop, detectable differences were evident in the 'Partial Condition' and '2 Valid Links,' but not with '1 Valid Link.' These findings suggest significant variations in the application of diverse search strategies for both current and voltage drop among the Control and Experimental group 2. In a comparison between Experimental groups 1 and 2, two variables stood out in post-test performance: "Post-test: Voltage Drop 1 Valid Link" and "Post-test: Current 2 Valid Links," indicating divergence in generating valid links. This reflects Novak and Treagust’s (2022) findings that well-constructed explanations foster connections between various scientific concepts. These post-test disparities reinforce Novak and Treagust's (2022) observation that students' comprehension and explanations tend to become more nuanced as they accrue experience. This evolution in understanding suggests that the educational tools used in the study offer diverse avenues for conceptual growth. Such variations in post-test outcomes are in alignment with the Knowledge Integration (KI) framework proposed by Linn and Eylon (2011), further emphasizing the role of these tools in facilitating a more sophisticated understanding of the subject matter.

In comparison to the Control group, Experimental groups 1 and 2 exhibited productive engineering approaches and search strategies, aligning with the idea that explanations evolve over time. This observation is congruent with the research conducted by Novak and Treagust (2022; 2018). According to McNeil and Krajcik (2011), an effective scientific explanation consists of three core elements: a claim, supporting evidence, and reasoning. Such an explanation not only connects diverse scientific concepts but also improves students' scientific understanding. As students gain experience and deepen their understanding, their explanations undergo a refinement process, approaching the expertise seen in practicing scientists. In this study, students in Experimental groups 1 and 2 were tasked with making predictions and providing the underlying rationale through the use of rule formation tools. They were also required to document their findings and reasoning post-experimentation. This iterative process of crafting explanations enforced students to engage more thoroughly with scientific concepts, echoing the findings of Fortus and Krajcik (2012) and Krajcik and Shin (2014). This type of evolving explanation, which incorporates more valid integrated ideas over time as students gain experience, can be seen in the statistically detectable differences between the groups in their posttest knowledge integration scores, as evaluated using the knowledge integration rubric. Novak and Treagust (2022) further emphasize the role of scaffolds in enabling students to undertake complex tasks (Quintana et al., 2004; Tabak, 2004; Wood, Bruner, & Ross, 1976; McNeil & Krajcik, 2011; Braaten & Windschitl, 2011). These findings align with these observations, where the Experimental 1 group, equipped with a decision table and induction rule tool, and the Experimental 2 group, provided only with the induction rule tool, demonstrated significant differences in multiple areas compared to the Control group. Novak further argued that constructing scientific explanations demands time, several exposures, and feedback (Fortus & Krajcik, 2012; Novak & Treagust, 2018). Although the tools used in this experiment did not offer direct feedback, they enabled students in the experimental groups to make meaningful connections. Specifically, they used the decision table and the rule induction tool to build an integrated understanding of electric circuit simulations, thereby enhancing their overall performance.

In summary, the observed differences between the Control and Experimental groups in this study highlight the influence of varying tools and approaches on multiple facets of learning, including prediction, confidence, and rule formation, all underpinned by extant theories and models in the field of educational psychology and science education. Novak and Treagust's (2022) suggestion that instructional materials should revisit previous ideas and introduce new ones (NASEM, 2019; Bransford et al., 2000; NRC, 2012) is mirrored in this study by the tools given to the Experimental 1 and 2 groups that prompted them to reevaluate their predictions and adjust their claims. The absence of direct feedback from the tools did not prevent students from considering new ideas, making connections between ideas, and adjusting their claims. The observed differences between the Control and the Experimental groups in various variables, including posttest knowledge integration scores, support the effectiveness of Novak and Treagust's (2022) approach in science education.

5.2. RQ 2. What distinct learning paths do students take as they engage with the electric circuit simulation in the control condition versus rule formation conditions?

In this sequence analysis, I strategically focus on overarching patterns of clusters observed across all participant groups — Control, Experimental 1, and Experimental 2 — as opposed to dissecting patterns specific to each group. This choice is anchored in several core considerations.

A primary rationale stems from the numerical composition of groups in the study. Each group consists of only 20 participants, and analyzing these clusters in isolation could risk undermining the robustness of findings due to this limited sample size. By unifying all groups for analysis, I amplified the sample size to a total of 60 participants. This consolidation enhances the reliability of findings and, because instructional circumstances also vary, broadens perspective and enabling extracting more comprehensive insights into the universal pattern of clusters.

The secondary impetus is derived from the common threads discernible across all groups despite their distinct cluster patterns. It is important to clarify that I am not asserting the Control group’s cluster patterns to be identical to those in the experimental groups. In fact, the clusters present diverse trends reflective of the unique interventions each group experienced. The Control group, equipped with a drawing tool, predominantly exhibited non-normative ideas. Conversely, Experimental groups 1 and 2, which received decision tables and induction rule tools, demonstrated partial and multiple valid links. For instance, findings (as illustrated in Figure 4.23) revealed that the search strategy patterns in the Control group manifested two distinct clusters. These clusters ultimately culminated in either non-normative ideas on the posttest or a combination of a partial link with non-normative ideas. In contrast, the Experimental 1 group's search strategy patterns yielded three clusters (Figure 4.21). These clusters led to a range of outcomes: partial links for both current and voltage drop; a combination of one valid link for voltage drop with multiple valid links for current; and finally, multiple valid links for both current and voltage drop. Meanwhile, in the Experimental 2 group, search strategy patterns resulted in three clusters (Figure 4.22). These clusters culminated in diverse outcomes, including partial links; a combination of a partial link for voltage drop with one valid link for current; and a blend of a partial link for current and multiple valid links for voltage drop. These patterns were similar to the five identified clusters across all groups, further justifying the decision to amalgamate.

Despite these individual differences, an amalgamated view allows us to identify shared outcomes across all groups. This approach facilitates the recognition of a comprehensive five-cluster structure that spans all participant groups (Figure 32). These clusters represent distinct sequences of search strategies and a spectrum of knowledge integration, from non-normative ideas, to partial links, to one partial link, and to multiple valid links. For example, when we examine all groups collectively, a cluster characterized by non-normative ideas, similar to the ones observed in the Control group, also emerges. Similarly, the experimental groups' patterns of partial and multiple valid links are echoed in the clusters identified across all groups. Upon a more holistic assessment, I identified a consistent pattern within the clusters across all groups. The first cluster was associated with non-normative ideas, the second exhibited partial links, the third revealed one partial link for voltage drop paired with one valid link for current, the fourth showed one partial link for voltage drop alongside multiple links for current, and the fifth demonstrated multiple valid links for both voltage drop and current.

Therefore, the rationale for analyzing the patterns of clusters among all groups, rather than within each individual group, lies in both statistical robustness provided by the larger sample size, and the relative consistency in the patterns of clusters found in each individual group as compared to those across all groups.

5.2.1. Patterns of Search Strategies among Clusters

The patterns discerned from the sequential use of search strategies across the five clusters reveal approaches taken by each cluster in scientific discovery. Beginning with the sequence of prediction, clusters 1 primarily relied on using the same rule for making predictions. This pattern is akin to the use of prior knowledge to predict outcomes, a strategy that aligns with the discovery learning principles proposed by Bruner (Bruner, 1961). Bruner emphasized the use of existing schemas to understand and predict new information, which is reflected in the prediction strategies used by clusters 1 and 2.

However, in contrast, clusters 2, 3, 4 and 5 used a different approach to prediction. They implemented a 'fill up gaps' strategy, indicating the students' efforts to identify and address their knowledge gaps. This pattern aligns with the concept of self-regulated learning, where learners actively engage in monitoring their learning process and identifying their learning needs (Zimmerman, 2000). The use of this strategy in clusters 3 and 5 shows a learner-centered approach that underscores self-regulation.

When looking at the confidence measures (verify vs. falsify), the clusters show a variety of approaches. In clusters 1, 2, and 3, the students primarily used verification to ascertain their predictions. This aligns with the discovery learning strategy where learners seek to confirm their understanding through exploration and testing (Bruner, 1961). In contrast, cluster 4 and 5 show an instance of falsifying predictions. This reflects a critical approach to learning, where learners question their assumptions and understanding - a key component of self-regulated learning (Zimmerman, 2000).

The rule formation strategies in the clusters showed significant variability. The ‘confirmation redundancy’ strategy used in clusters 1 and 3 indicates an emphasis on reinforcement learning where learners repeatedly test the same instance to consolidate their understanding. This rule aligns with the early stages of knowledge integration, where learners often revert to non-normative or partially linked ideas. By repeatedly testing the same instance, learners are attempting to confirm and reconfirm their initial understanding of the phenomena. While this strategy is conducive to reinforcing specific understandings, it may limit learners' ability to progress towards generating multiple valid links, as it restricts the exploration of new hypotheses and possibilities. However, repeated testing of the same instance could also signify the consolidation of a particular link, potentially transitioning from a partial link to a valid link. On the other hand, the 'successive scanning' and 'simultaneous scanning' strategies used in clusters 2, 3, 4, and 5 align with the heuristic approach where learners use simplified strategies or "rules of thumb" to understand complex systems (Shah & Oppenheimer, 2008). These rules indicate the learners' cognitive flexibility in understanding complex systems, employed in Clusters 2, 3, 4, and 5. With successive scanning, learners evaluate one attribute at a time, aligning with the formation of a single valid link in the knowledge integration process. This strategy allows learners to investigate one aspect of the phenomena and move onto the next, promoting disconnected discoveries. In contrast, simultaneous scanning requires the learners to evaluate all attributes concurrently and eliminate attributes they feel confident about, limiting the formation of multiple valid links when not enough explorations are made. The 'focus gambling' found in Cluster 3, signifies the learners' propensity to take calculated risks based on their current understanding, indicating more intense self-regulated learning. Learners might make assumptions or predictions that could lead to a breakthrough in their understanding, possibly transitioning from a partial link or a single valid link to multiple valid links. However, the effectiveness of this strategy would largely depend on the learner's current understanding, ability to handle failure and adjust their strategies accordingly, and willingness to confirm their discoveries with further testing. Moreover, the 'conservative focusing' observed in clusters 4 and 5 is indicative of learners' systematic approach to problem-solving, further supporting the application of self-regulated learning strategies. Learners applying this rule only change one attribute at each trial, allowing for careful observation of cause and effect. This aligns with the process of formulating valid links in the knowledge integration framework, as it promotes an iterative, step-by-step understanding of the system. Conservative focusing could be particularly useful in situations where learners have established partial links but are struggling to transition towards multiple valid links, as it provides a clear and focused pathway for deeper exploration.

The sequential analysis yielded five distinct clusters, representing different approaches and strategies students employed throughout the learning process. It is apparent that students adapted different engineering approaches and search strategies as they progressed through the simulations. Across all five clusters, students demonstrated various stages and strategies of Bruner's discovery learning, utilizing engineering approaches and search strategies to test, predict, falsify and/or verify predictions after testing, and formulate rules within the electric circuit simulation. Furthermore, the sequence analysis provides insight into the students' self-regulated learning processes, where the students not only learned about the content but also modified their learning strategies as they engaged with the simulation. The notion of deliberate practice in developing self-regulated learning skills, as discussed by Winne (2022) and supported by research (Ericsson & Harwell, 2019), is evident in the iterative nature of the students' learning observed in the clusters. The students engaged in deliberate practice by actively making predictions, verifying them, and adjusting their strategies based on the outcomes they encountered during the electric circuit simulation. This iterative process reflects deliberate practice and adaptive learning, as the students continuously refined their understanding and rule formulation strategies.

5.2.2. Patterns of Search Strategies in Each Cluster

This study identified different learner clusters, each demonstrating unique cognitive models and learning paths. This aligns with Obaid et al.'s (2023b) findings, as they also identified varied learning paths among students based on their instructional conditions. The variations in learners' strategies and the influence of instructional conditions on learning paths are common findings across both studies. Findings and the results of Obaid et al. (2023b) contribute to the broader understanding of scientific discovery learning and knowledge integration in instructional conditions. The primary emphasis in both studies is on understanding students' learning paths, the strategies they employ in simulation-based discovery learning, and the importance of knowledge integration and self-regulated learning.

5.2.2.1 Reinforced Confirmers

This group heavily relied on the 'confirmation redundancy' rule, endeavoring to validate and bolster their pre-existing comprehension, without venturing into unknown territories to generate and implement new rules. In Cluster 1, the learners made predictions and sought verification using established rules. However, no evidence of problem identification surfaced. This suggested a type of learner who readily applied known rules but grappled with recognizing issues and creating novel rules, a pattern similarly observed in certain educational research (De Jong & van Joolingen, 1998; Klahr & Nigam, 2004). They exhibited a preference for the ‘confirmation redundancy’ rule, which according to Klahr, Chen, & Toth (2001), mirrors an initial approach in scientific discovery learning. The students partook in confirming redundancy, demonstrating an eagerness to confirm and fortify their initial understandings via repeated testing. Yet, the variance in post-test outcomes suggested this methodology might have confined their exposure to alternate perspectives or strategies. For Cluster 1, the persistent dependence on the ‘confirmation redundancy’ rule indicated a certain level of comfort in relying on previously established norms. This, however, also hinted at a deficiency in exploration, possibly stemming from a higher requirement for cognitive certainty, a characteristic tied to a lower tolerance for ambiguity and uncertainty (Furnham & Marks, 2013). For these students, instructional cues encouraging exploration and explication of the advantages of making and learning from errors could prove beneficial. Bruner's notion of discovery learning underlines the significance of exploring diverse approaches, which seemed to be somewhat lacking in these clusters. According to Winne (2022), effective learners often engage in multiple phases of self-regulated learning, which includes goal setting, planning, and adapting strategies. The "Reinforced Confirmers", however, seem largely confined to the 'confirmation redundancy' strategy. They appear to lack extensive goal-setting and planning, given that they don't venture into the unknown and remain largely fixated on confirming existing understandings. In terms of monitoring, they are only concerned with confirming their pre-existing beliefs, thus lacking the adaptive element. Their approach aligns with Winne's notion of a more rudimentary form of self-regulated learning that does not fully exploit the adaptive or monitoring components. As a result, educational strategies should aim to urge students to explore a multitude of approaches.

5.2.2.2 Dual-Mode Strategy Diversifiers

This group exhibited adaptability by effectively implementing both 'successive scanning' and 'simultaneous scanning' strategies in formulating rules. This indicates a versatility in their thought process and learning pathways. In Cluster 2, learners made predictions and verified these using established rules, yet, once again, significant problem identification was conspicuously absent. However, this cluster exhibited a notable distinction in its approach to rule formulation for current and voltage drop, demonstrating flexibility in utilizing different strategies of 'successive scanning' and 'simultaneous scanning.' This aligns with studies (Kuhn, Garcia-Mila, Zohar, & Andersen, 1995; Schauble, 1990) that propose that multiple strategies can coexist within the same individual, emphasizing the need for flexibility in learners' thinking and approaches. Similarly, this aligns with Obaid et al.'s (2023b) finding that different students have different learning paths that require specific prompts and strategies to aid their understanding. They observed that these learning paths were contingent upon instructional conditions, a finding echoed by the varying learning paths evident in the clusters of the current study based on different rule formulations. The "Dual-mode Strategy Diversifiers" are more adaptable in their learning paths. According to Winne's (2022) framework, this demonstrates an ability for more complex planning and monitoring. However, this cluster still lacks in problem identification, a key element of Winne's adaptive and monitoring phases. This suggests that while they may set initial goals and plans, their lack of problem identification could indicate a deficit in the monitoring and adapting phases of self-regulated learning. In these instances, prompts that encourage students to reflect on the problems they encounter or their learning process, as suggested by Obaid et al. (2023b) and Panadero (2017), might prove beneficial.

5.2.2.3 Multi-Strategy Jugglers

This group implemented an array of strategies, showcasing a more intricate and dynamic cognitive model that traverses various strategic paths. Cluster 3 demonstrated a more dynamic methodology, employing a plethora of rule generation strategies. The group manifested the use of diverse rule formation strategies, including 'successive scanning,' 'simultaneous scanning,' and 'focus gambling,' and managed to attain a level of knowledge integration by the conclusion of the process, indicating a more sophisticated cognitive model. This concurs with Winne's (2018) research, suggesting learners who can flexibly apply a variety of strategies are more likely to attain learning goals. Moreover, this cluster emphasizes the significance of self-regulated learning, consistent with Zimmerman's (2002) theory that self-regulation plays a pivotal role in the process of learning and cognitive development. This notion resonates with Bruner's principles, as the students were not merely passively receiving information but were actively participating in the learning process. In contrast to the "Reinforced Confirmers" and "Dual-Mode Strategy Diversifiers," the Multi-Strategy Jugglers exhibited a superior level of knowledge integration in their post-test assessments. This is likely attributable to their expansive strategic repertoire and supports the knowledge integration framework, which argues that the interweaving of various ideas amplifies understanding. However, it's worth noting that while this group displayed characteristics of adaptive learning, the specific strategies they employed may not be the most efficient for complex problem-solving or achieving higher levels of knowledge integration. In other words, the mere possession of multiple strategies is insufficient; the efficacy of those strategies is equally pivotal. As I look to the future, instructional designs should capitalize on this nuanced, integrated approach, especially in the context of simulation-based learning environments.

5.2.2.4 Self-Regulated Revisers

They showed significant elements of self-regulated learning by adjusting their strategy based on their success in predicting results. They recognized when a change in strategy was needed, emphasizing their adaptive nature. Cluster 4 displayed flexibility and self-regulation in their learning process. They showed an inclination to reassess their understanding, as observed through the falsification of voltage drop prediction. This critical thinking element aligns with the tenets of self-regulated learning, wherein students actively control their learning processes, including cognitive strategies. In Cluster 4, the learners demonstrated an understanding of the importance of making predictions, verifying these, and identifying issues when inconsistencies arise. This aligns with the notion of self-regulated learning as proposed by Pintrich and De Groot (1990), which suggests that self-regulating learners are more proficient at setting goals, monitoring progress, and adjusting strategies as needed. Notably, this cluster also showed the ability to falsify a prediction, underscoring an important aspect of self-regulated learning – the ability to recognize when a change in strategy is necessary (Schraw, Crippen, &

Hartley, 2006). Tools such as decision tables could have been beneficial here to systematize their rule generation process.

5.2.2.5 Methodical Integrators

These students consistently verified or falsified their predictions and switched to more focused and methodical approaches like 'conservative focusing' in their rule formation, enabling them to create multiple valid links and integrate knowledge. Cluster 5 demonstrated an adaptive and methodical approach to learning, characterized by consistent verification or falsification of their predictions. In their rule formation, they adopted a more systematic and focused approach, switching to conservative focusing. The consistent application of different rule-generation strategies for current and voltage drop indicates a thoughtful selection process, potentially enabled by the decision table and rule induction tool. This behavior exemplifies effective self-regulated learning, leading to multiple valid links between different concepts, as indicated by their post-test results. This aligns with research by Linn et al. (2004) who argue that learners must be able to integrate and link knowledge to fully understand scientific concepts. The "Methodical Integrators" seem to be the epitome of Winne's (2022) self-regulated learners. They engage in goal setting by filling up gaps in their predictions, indulge in planning through multiple rule-generation strategies, and constantly monitor and adapt through verification or falsification of predictions. Moreover, their use of conservative focusing alongside other strategies aligns with Winne's notion that effective strategy regulation involves choosing the most effective strategies for a given context. They are not just using multiple strategies; they are selectively applying the best strategies, resulting in more knowledge integration in the posttest. This suggests that providing students with structured tools to guide their learning can facilitate knowledge integration.

In this study, knowledge integration, characterized by the interlinking of ideas from formulated rules, was found to enhance understanding, particularly in Clusters 3, 4, and 5. This is congruent with Obaid et al.'s (2023b) findings. They noted that when students successfully distinguish ideas and form a coherent integration, they exhibited better knowledge integration. Furthermore, both this study and Obaid et al.'s underline the importance of self-regulated learning. This study's findings for Clusters 3, 4, and 5, which exhibit flexibility and methodical learning approaches, align well with Obaid et al.'s emphasis on self-regulated learning. This suggests a consistency in the belief that self-regulated learning fosters better knowledge integration.

The study supports previous findings that learners often employ a positive test strategy to validate their hypotheses with fitting cases rather than non-fitting ones, as originally identified by Wason (1960) and later interpreted by Klayman and Ha (1987). This pattern was particularly observed in the "Reinforced Confirmers" cluster, where students exhibited a strong tendency toward confirmation redundancy, seeking to validate and bolster their pre-existing comprehension without venturing into uncharted territories to generate and implement new rules. It is worth noting, however, that utilizing a positive test strategy should not be seen as a lack of hypothesis-testing skills or an unwillingness to disprove hypotheses. Participants often engage in "limit testing," trying to maximize the chance of falsification within a positive testing framework.

The tendencies and behaviors exhibited by this study's identified clusters reinforce the observation that learners often make overly restricted hypotheses, favoring conditions that are sufficient but not necessarily required. The most glaring example comes from the "Reinforced Confirmers" cluster. This group of learners favored positive testing, heavily relying on the 'confirmation redundancy' rule to fortify their pre-existing understanding. In other words, they were primarily concerned with validating their current knowledge rather than venturing into unexplored territories or testing their knowledge against counter-evidence. This behavior serves as a textbook example of learners creating overly restricted hypotheses, emphasizing conditions they perceived as sufficient, but which may not have been necessary.

Similarly, the "Dual-mode strategy diversifiers" cluster, despite demonstrating versatility in their thought processes and rule formulation strategies, did not veer too far from the path of positive testing. They adeptly employed 'successive scanning' and 'simultaneous scanning' strategies, but significant problem identification was notably absent. This suggests that while these learners were comfortable making predictions and verifying them using established rules, they showed little inclination to actively disprove their hypotheses or seek alternative perspectives.

On the other hand, the clusters of "Self-regulated revisers" and "Methodical integrators" demonstrated traits of negative hypothesis testing. The "Self-regulated revisers" displayed an inclination to reassess their understanding, as evident through the falsification of voltage drop predictions. Their learning approach indicated an active engagement in identifying inconsistencies and adjusting strategies based on predictive success - the core tenets of negative testing. The "Methodical integrators" also presented a blend of both positive and negative testing. They methodically and consistently verified or falsified their predictions, showcasing a willingness to challenge their hypotheses, which led to the creation of multiple valid links and effective knowledge integration.

These behaviors underscore the ecological basis for learners favoring positive testing, as Klayman & Ha (1987) theorized. In particular, the "Reinforced Confirmers," with their evident discomfort venturing outside of established knowledge, possibly perceive false positives as more threatening, potentially destabilizing their cognitive certainties. This kind of learner, marked by a high requirement for cognitive certainty and a low tolerance for ambiguity and uncertainty, may view the risk of false positives as more harmful than that of false negatives, contributing to their favoring of positive hypothesis testing.

However, successful hypothesis testing is not merely about confirmation or disconfirmation. Instead, as highlighted by Platt (1964) and Wason & Johnson-Laird (1972), it depends on the effective use of alternative hypotheses. This is something I observed in the "Multi-strategy jugglers" and "Self-regulated revisers" clusters. These participants demonstrated a higher tendency to test alternatives, doing so in almost one out of three trials. The less successful subjects did so in about one out of nine trials. The use of alternative hypotheses guided negative tests, indicating plausible target instances outside the current hypothesis. They also guided double negative tests by identifying non-target instances within the current hypothesis. Testing alternative hypotheses contributed to the discovery process as the falsification of a current hypothesis established a new one for continued investigation. Without alternatives, subjects may struggle to learn from falsifications.

Hypothesis revision, as emphasized by Lakatos (1970, 1978), is as vital as hypothesis testing in rule discovery tasks. This principle was clearly demonstrated in the "Self-regulated revisers" and "Methodical integrators" clusters, who showed a strong tendency to adjust their strategies based on the success or failure of their predictions. Even if subjects don't voluntarily generate alternatives, they are forced to do so when their current hypothesis is disproved. This highlights the importance of fostering an environment that encourages learners to embrace falsification and regard it as an opportunity to adjust and enhance their understanding.

Future research should focus on how feedback informs hypothesis revisions, particularly considering that hypothesis testing is a process involving the identification of relevant features and continuous refinement of one's conceptual model of the rule or phenomenon.

5.3. Implications of Cluster Analysis for Instructional Interventions

The identification of distinct clusters based on learning trajectories offers insights for tailoring instructional interventions in simulation-based discovery learning. Each cluster's unique approach to prediction, confidence measures, problem identification, rule formation, and knowledge integration suggests specific areas where learners can benefit from targeted support.

For example, learners in Cluster 1, characterized by heavy reliance on existing knowledge without much exploratory behavior, would be predicted to benefit from prompts encouraging alternative hypotheses and experimentation. Implementing prompts like "What alternative explanations could account for your observations?" and "How might changing the resistance impact the circuit?" could facilitate deeper engagement with the exploratory aspect of the learning process.

Cluster 2 learners, displaying an iterative approach, presumably would benefit from scaffolding that prompts reflection on the connections between observations and predictions. Questions like "How does your observation align with or challenge your initial prediction?" and "What new insights can you derive from your latest trial?" could foster better integration of knowledge.

Cluster 3 students, who maintain a consistent prediction rule while varying rule-generation strategies, likely would find value in prompts that help identify gaps in understanding and encourage comprehensive rule formulation. Suggestions like

"Compare your current prediction to previous ones" and "What patterns can you identify to formulate a more comprehensive rule?" could facilitate better rule refinement.

Similarly, for Cluster 4 learners who exhibit flexibility in learning strategies, prompts could encourage critical reflection on strategy effectiveness. Questions such as "Under what conditions would you revise your prediction?" and "Can you design an experiment to challenge one of your rules?" may further enhance their learning process.

Lastly, learners in Cluster 5, demonstrating systematic and focused behavior, could benefit from prompts that encourage deeper exploration and application of findings. Prompts like "What broader implications do your findings suggest?" and "How can you apply your learning to a different circuit?" could encourage extended thinking.

5.4. Implications for Learning and Practice

The analysis of distinct clusters based on learning trajectories not only illuminates effects of tailored instructional interventions but also offers insights for pedagogical practices in the realm of simulation-based discovery learning.

The existence of distinct clusters underscores the notion that learners engage with simulation-based discovery learning in varied ways. This has a significant and oft-noted implication for educational practice; there isn't a "one-size-fits-all" approach. Students might benefit from a more personalized learning environment where their unique learning paths are recognized and nurtured.

Insights from this study can be incorporated into curriculum development, offering educators a roadmap for weaving in targeted prompts and reflection questions throughout simulation-based discovery learning modules. This can ensure consistent reinforcement and support for students, aligning with their identified learning trajectories.

Exploration and reflection emerge as key components. Encouraging students to actively engage with learning material and reflect on their experiences can foster deeper understanding. This suggests that instructional environments should embed opportunities for students to hypothesize, test, and rethink their ideas.

The clusters indicate varying degrees of metacognitive awareness and application in simulation-based discovery learning. Metacognitive skills, including the ability to evaluate one's own understanding, adjust strategies when necessary, and predict outcomes, are critical for effective discovery learning. Fostering these skills might enhance students' learning experiences and outcomes.

The findings of this research have implications for the development of educational resources and tools, particularly simulation-based tools. Developers should aim to create resources that are aligned with the principles of discovery learning and facilitate the development of effective search strategies and heuristics for diverse learning paths. Developers can use insights developed by examining different learning clusters to create more adaptive learning technologies that are capable of identifying learners' tendencies and providing real-time, tailored support. This should not only enhance the effectiveness of technology-enhanced learning environments but also improve alignment of tools with learners' natural learning processes.

5.4. Limitations and Future Research Directions

It is important to acknowledge the limitations of the study. Although statistically detectable differences were observed between the groups, caution should be exercised when interpreting these results. Due to the complex nature of learning and the multitude of factors influencing it, tools provided in this research are not the only determinant of learning success. Other individual and contextual factors, such as motivation and learning environment, can also significantly influence learning outcomes. Furthermore, the study was conducted within the context of learning about electric circuits. The generalizability of these findings to other domains within science education or other fields requires further investigation.

In future research, it would be valuable to conduct similar studies with a larger sample size, or to utilize a mixed-methods approach to further understand how learners interact with the simulation environment. Moreover, investigating other variables, such as learner satisfaction and long-term knowledge retention, would add further depth to the understanding of simulation-based discovery learning.

Future research could investigate the optimization of these tools, by perhaps integrating adaptive features that align with the learner's pace and understanding. Additionally, exploring the interplay of individual differences with the tools provided could yield insights into how best to support diverse learners in simulation-based environments. The role of feedback within these environments, and how it can be employed to further support self-regulated learning and knowledge construction, is another avenue worthy of exploration. Based on the prototype sequences identified in this study, future research should explore the efficacy of adaptive AI tutoring dialogues that can provide personalized guidance and support to students based on their current stage in the discovery learning process. The identified patterns of sequence in each cluster can serve as student models for future research, particularly for creating adaptive AI tutoring dialogues. For instance, Cluster 1's model could benefit from AI-driven prompts encouraging problem identification and metacognitive thinking. Similarly, Cluster 2's model could benefit from prompts focused on critical reflection on the effectiveness of their iterative rule adjustments. Cluster 3's model could leverage AI prompts that stimulate further variety in rule formulation and foster problem identification. Cluster 4's model, with its high degree of self-regulation, might best be served with AI prompts that encourage critical reflection and strategy refinement. Lastly, Cluster 5's model, with its balance of verification, falsification, problem identification, and shifting strategies, would benefit from prompts that stimulate further exploration, questioning, and discovery.

These findings have significant implications for educators and software developers. For educators, they underline the importance of providing students with tools that can guide them in self-regulated learning and knowledge integration. For software developers, these results suggest that the inclusion of decision-making tools and rule induction tools could enhance the educational effectiveness of simulation-based learning environments.

5.4. Contribution to Educational Practice and Scholarly Significance

These findings hold significance for both educators and software developers. Educators can leverage tools to guide students in self-regulated learning and knowledge integration, while software developers can enhance educational efficacy through decision-making and rule induction tools. The study's theoretical implications aligning with Bruner's discovery learning, self-regulated learning, and the Knowledge Integration Framework underscore the need for multifaceted approaches in designing effective learning environments.

First, the findings contribute to the field of education and cognitive psychology in several ways. The distinct sequential patterns exhibited by each cluster reflect different learning processes at play in simulation-based discovery learning. These processes correspond to the different cognitive strategies proposed by Bruner's discovery learning theory, such as simultaneous scanning, successive scanning, confirmation redundancy, focus gambling, and conservative focusing. By analyzing how these strategies unfold in sequence in each cluster, we gain a detailed understanding of how students navigate and learn from simulation-based environments.

Second, the study provides insights into self-regulated learning in simulation-based learning environments. The patterns of prediction, problem identification, rule formation, and post-test knowledge integration demonstrate how students monitor and adjust their learning strategies. For instance, in Cluster 4, the students exhibited flexibility and self-regulation by switching their rule formation strategies across trials, and even falsifying their predictions when necessary. Such behaviors indicate active engagement in self-regulated learning.

Third, the study's findings shed light on the knowledge integration process in simulation-based learning. By analyzing the sequences of prediction, confidence measures, problem identification, rule formation, and knowledge integration, we can identify the pathways through which students build, revise, and integrate their understanding of the concepts. For example, in Cluster 5, the students demonstrated progressive learning and successful knowledge integration as they adapted their prediction and rule formation strategies over the trials, resulting in the formation of multiple valid links by the post-test.

This study contributes to the understanding of how students engage with simulation-based discovery learning, particularly in forming predictions, employing confidence measures, and generating rules. By linking students’ strategies and approaches to Bruner's discovery learning, this study provides insight into how classic cognitive psychology theories can still be applicable in contemporary learning environments. Moreover, the study contributes to the development of student models for creating adaptive AI tutoring dialogues. The distinct sequential patterns in each cluster can serve as prototypes for different types of learners. For example, a student model based on Cluster 1 might represent learners who heavily rely on existing knowledge without much exploratory behavior, while a student model based on Cluster 4 might represent learners who exhibit high flexibility and self-regulation. Such student models can inform the design of personalized instructional interventions in AI tutoring systems.

Overall, this study contributes to understanding how various tools can facilitate learning in simulation-based environments. It also highlights the value of combining principles from Bruner's discovery learning strategies, self-regulated learning, the Knowledge Integration Framework, and cognitive psychology in designing effective learning environments. By applying these principles and making use of appropriate tools, learners can be better supported in their journey of discovery and understanding.

In light of these results, scaffolding might be further explored to support the discovery process in educational contexts. Instructional interventions, such as decision tables and rule induction tools provided to experimental groups, can foster students' abilities to develop and refine their rules, as shown by Quintana et al., (2004). Furthermore, these findings also point towards the potential of utilizing the student models derived from these clusters to create adaptive AI tutoring dialogues. This adaptive instruction can foster deeper learning by dynamically adjusting to learners' needs and promoting self-regulated learning.

In conclusion, this study enriches understanding of diverse learning trajectories in simulation-based discovery learning. Through the identification of distinct clusters characterized by different patterns of prediction, confidence measures, problem identification, rule formation, and knowledge integration, the study offers valuable insights that can inform educational interventions and the development of adaptive learning technologies. The integration of theoretical perspectives, including Bruner's discovery learning, self-regulated learning, and the knowledge integration framework, provides a multifaceted lens through which the complexities of learning in simulation-based environments can be better understood and supported.

References

Ainsworth, S. (1999). The functions of multiple representations. *Computers and Education*, *33*(2), 131–152. https://doi.org/10.1016/S0360-1315(99)00029-9

Alfieri, L., Brooks, P. J., Aldrich, N. J., & Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology, 103*(1), 1–18. https://doi.org/10.1037/a0021017

Bandura, A. (1997). *Self-efficacy: The exercise of control*. W H Freeman/Times Books/ Henry Holt & Co.

Bies, B., Dabbs, K., Zou H., 2006. On determining the number of clusters: A comparative study. http://www.ima.umn.edu/preprints/aug2009/2272.pdf

Braaten, M., & Windschitl, M. (2011). Working toward a stronger conceptualization of scientific explanation for science education*. Science education, 95*(4), 639– 669. https://doi.org/10.1002/sce.20449.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (1999). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Research Council.

Bransford, J.D., & Schwartz, D.L. (1999). Rethinking transfer: A simple proposal with multiple implications*. Review of Research in Education, 24*, 61-100. https://doi.org/10.2307/1167267.

Brockbank, E., & Walker, C. M. (2022). Explanation impacts hypothesis generation, but not evaluation, during learning. *Cognition*, *225*, 105100–105100. https://doi.org/10.1016/j.cognition.2022.105100

Brockbank, E., Lombrozo, T., Gopnik, A., & Walker, C. M. (2023). *Ask me why, don't tell me why: Asking children for explanations facilitates relational thinking*.

Bruner, J.S., Goodnow, J.J., & Austin, G.A. (1956). *A study of thinking*. New York: NY Science Editions, Inc.

Chambers, S. K., C. Haselhuhn, T. Andre, C. Mayberry, S. Wellington, A. Krafka, J. Volmer, and J. Berger. (1994). The Acquisition of a Scientific Understanding of Electricity: Hands-On versus Computer Simulation Experience; Conceptual Change versus Didactic Text. Paper presented at the Annual Meeting of the American Educational Research Association 1994, New Orleans, LA. doi: 10.3168/jds.S0022-0302(94)77044-2

Correia, A.-P., Koehler, N., Thompson, A., & Phye, G. (2019). The application of PhET simulation to teach gas behavior on the submicroscopic level: secondary school students' perceptions. *Research in Science & Technological Education*, *37*(2), 193–217. https://doi.org/10.1080/02635143.2018.1487834

de Jong, T. & van Joolingen, W. R. (1998) Scientific discovery learning with computer simulations of conceptual domains. Review of Educational Research, 68, 179-

201. https://doi.org/10.3102/00346543068002179

Dewey, J. (1910). *How we think*. D C Heath. https://doi.org/10.1037/10903-000
Duffy, T. M., and D. J. Cunningham. (1996). Constructivism: Implications for the Design and Delivery of Instruction. New York: Simon & Schuster Macmillan.

Ericsson, K. A., & Harwell, K. W. (2019). Deliberate practice and proposed limits on the effects of prac- tice on the acquisition of expert performance: Why the original definition matters and recommen- dations for future research. Frontiers in Psychology, 10. https://doi.org/10.3389/fpsyg.2019.02396

Farris, H. H., & Revlin, R. (1989). Sensible reasoning in two tasks: Rule discovery and hypothesis evaluation. *Memory & cognition*, *17*, 221-232.

Fortus, D., & Krajcik, J. S. (2012). In Curriculum coherence and learning progressions, B. Fraser, K. Tobin, & C. McRobbie (Eds.), *Second international handbook of science education*, (vol. 52, pp. 783–798). Dordrecht: Springer. https://doi.org/10.1007/978-1-4020-9041-7.

Friedler, Y., Nachmias, R., & Linn, M.C. (1990). Learning scientific reasoning skills in microcomputer-based laboratories. *Journal of Research in Science Teaching, 27*, 173-191

Furnham, A. & Marks, J. (2013). Tolerance of Ambiguity: A Review of the Recent Literature. Psychology, 4, 717-728. doi: 10.4236/psych.2013.49102.

Géron, A. (2019). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow : concepts, tools, and techniques to build intelligent systems / Aurélien Géron.* (Second edition.). O'Reilly Media, Inc.

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: a response to Kirschner, Sweller, and Clark (2006). *Educational psychologist, 42*(2), 99-107.

Hodson, D. (1998). Science fiction: The continuing misrepresentation of science in the school curriculum. *Curriculum studies, 6*(2), 191-216.

Kalyuga, S., and J. Plass. (2007). Managing Cognitive Load in Instructional Simulations. In Proceedings of the IADIS International Conference E-Learning 2007, edited by M. B. Nunes and M. McPherson, volume 1, 27–34). IADIS Press.

Kalyuga, S. (2011). Informing: A Cognitive Load Perspective. Informing Science 14 (1): 33–45. doi:10.28945/1349.

Khan, S. (2011). New Pedagogies on Teaching Science with Computer Simulations. *Journal of Science Education and Technology 20* (3): 215–232. doi:10.1007/s10956-010-9247-2.

Kiayman, J., & Ha, Y.-W. (1987). Confirmation, disconfirmation,and information in hypothesis testing. *Psychological Review, 94*,211-228

Kim, B., & Reeves, T. C. (2007). Reframing research on learning with technology: in search of the meaning of cognitive tools. *Instructional Science*, *35*(3), 207–256. https://doi.org/10.1007/s11251-006-9005-2

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75–86.http://dx.doi.org.proxy.lib.sfu.ca/10.1207/s15326985ep4102_1

Klahr, D., Chen, Z., & Toth, E. E. (2001). Cognitive development and science education: Ships that pass in the night or beacons of mutual illumination? In S. M. Carver & D. Klahr (Eds.), Cognition and instruction: Twenty-five years of progress (pp. 75–119). Lawrence Erlbaum Associates Publishers.

Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science, 12*(1), 1-48. https://doi.org/10.1016/0364-0213(88)90007-9

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.

Kohonen, T. (1985). Median strings. *Pattern Recognition Letters 3*, 309–313. https://doi.org/10.1016/0167-8655(85)90061-3

Krajcik, J. S., Blumenfeld, P., Marx, R. W., & Soloway, E. (2000). Instructional, curricular, and technological supports for inquiry in science classrooms. In J. Minstrel & E. H. Zee (Eds.), *Inquiring into inquiry learning and teaching in science* (pp. 283-315). American Association for the Advancement of Science.

Krajcik, J., & Shin, N. (2014). Project-based learning. In R. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (Cambridge Handbooks in Psychology, pp. 275-297). Cambridge: Cambridge University Press. https://doi. org/10.1017/CBO9781139519526.018.

Kuhn, D., Garcia-Mila, M., Zohar, A., & Andersen, C. (1995). Strategies of knowledge acquisition. Monographs of the Society for Research in Child Development, 60(4), v–128. https://doi.org/10.2307/1166059

Kuhn, D., Black, J., Keselman, A., & Kaplan, D. (2000). The development of cognitive skills to support inquiry learning. *Cognition and Instruction, 18*(4), 495-523.

Lajoie, S. P., & Derry, S. J. (Eds.). (1993). *Computers as cognitive tools*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Lakatos, I. (1970). Falsification and the methodology of scientific research programmes. In I. Lakatos &. A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 91-105). London: Cam- bridge University Press.

Lakatos, I. (1978). *The methodology of scientific research programmes.* London: Cambridge University Press.

Langley, P., Simon, H. A., Bradshaw, G. L., & Zytkow, J. M. (1987). *Scientific discovery.*

Computational explorations of the creative processes. Cambridge, MA: MIT Press.

Larkin, J. H., & Simon, H. A. (1987). Why a Diagram is (Sometimes) Worth Ten Thousand Words. *Cognitive Science*, *11*(1), 65–100. https://doi.org/10.1016/S0364-0213(87)80026-5

Lazonder, A. W., & Harmsen, R. (2016). Meta-Analysis of inquiry-based learning: Effects of guidance. *Review of Educational Research, 86*(3), 681–718.

Lee, H., J. L. Plass, and B. D. Homer. (2006). Optimizing Cognitive Load for Learning from ComputerBased Science Simulations. Journal of Educational Psychology 98 (4): 902–913. doi:10.1037/ 0022-0663.98.4.902.

Leemkuil, H., T. de Jong, R. de Hoog, and N. Christoph. (2003). KM QUEST: A Collaborative Internet-Based Simulation Game. Simulation & Gaming 31 (1): 89–111. doi:10.1177/1046878102250605.

Leutner, D. (1993). Guided discovery learning with computer-based simulation games: Effects of adaptive and non-adaptive instructional support. Learning and Instruction, 3(2), 113–132. https://doi.org/10.1016/0959-4752(93)90011-N

Lewis, E.L., Stern, J.L., & Linn, M.C. (1993). The effect of computer simulations on introductory thermodynamics understanding. *Educational Technology, 33*, 45-58.

Linn, M. C., Davis, E. A., & Eylon, B.-S. (2004). The scaffolded knowledge integration framework for instruction. In M. C. Linn, E. A. Davis, & P. Bell (Eds.), I*nternet environments for science education* (pp. 47–72). Lawrence Erlbaum Associates Publishers.

Linn MC, Eylon B-S. (2011). Science learning and instruction: Taking advantage of technology to promote knowledge integration. Routledge.

Linn, M. C., & Eylon, B.-S. (2006). Science education: Integrating views of learning and instruction. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (2d ed. pp. 511–544). Mahwah, NJ: Erlbaum.

Linn, M. C., Lee, H.-S., Tinker, R., Husic, F., & Chiu, J. L. (2006). Teaching and assessing knowledge integration. *Science, 313*, 1049–1050.

Lowe, R. K. (1999). Extracting information from an animation during complex visual learning. *European Journal of Psychology of Education*, *14*(2), 225–244. https://doi.org/10.1007/BF03172967

Ma, J., and J. V. Nickerson. (2006). Hands-On, Simulated and Remote Laboratories: A Comparative Literature Review. ACM Computing Surveys 38 (3): 1–24. doi:10.1145/1132960.1132961.

MacGregor, J. N., & Cunningham, J. B. (2008). Rebus puzzles as insight problems. *Behavior Research Methods, 40*(1), 263–268. https://doi.org/10.3758/BRM.40.1.263

Mayer, R. E. (1997). Multimedia learning: Are we asking the right questions? *Educational*

Psychologist, *32*(1), 1–19. https://doi.org/10.1207/s15326985ep3201_1

Mayer, R. E. (2009). Multimedia learning / Richard E. Mayer. (Second edition.). Cambridge University Press.

Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction. American Psychologist, 59(1), 14–19.

McNeill, K. L., & Krajcik, J. (2011). *Supporting grade 5–8 students in constructing explanations in science: The claim, evidence and reasoning framework for talk and writing*. New York, NY: Pearson Allyn & Bacon.

Minstrell, J. (2000). Student thinking and related instruction: Creating a facet-based learning environment. In J. Pellegrino, L. Jones, & K. Mitchell (Eds.) *Grading the Nation's Report Card: Research for the Evaluation of NAEP*. Committee on the Evaluation of NAEP, Board on Testing and Assessment. Washington, DC: National Academy Press.

National Academies of Sciences, Engineering, and Medicine (2019). Sc*ience and engineering for grades 6–12: Investigation and design at the center*. Washington, DC: The National Academies Press. https://doi.org/10.17226/2 5216.

National Research Council (2012). In H. Quinn, H. Schweingruber, & T. Keller (Eds.), *A Framework for K-12 science education: Practices, crosscutting concepts, and core idea*s. Washington, DC: The National Academies Press.

Navarro, G. (2001). A guide tour to approximate string matching. ACM Computing Surveys 33 (1), 31-88. https://doi.org/10.1145/375360.375365

NGSS Lead States. (2013). Next generation science standards: For states, by states. Washington, DC: The National Academy Press.

Nisbett, R., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.

Njoo, M., & De Jong, T. (1993). Exploratory learning with a computer simulation for control theory: Learning processes and instructional support. *Journal of Research inScience Teaching, 30*(8), 821-844. doi:10.1002/tea.3660300803

Novak, A. M., & Treagust, D. F. (2018). Adjusting claims as new evidence emerges: Do students incorporate new evidence into their scientific explanations? *Journal of Research in Science Teaching, 55*(4), 526–549. https://doi.org/10.1 002/tea.21429.

Novak, A. M., & Treagust, D. F. (2022). Supporting the development of scientific understanding when constructing an evolving explanation. *Disciplinary and Interdisciplinary Science Education Research*, *4*(1), 1–22. https://doi.org/10.1186/s43031-021-00043-w.

Obaid, T., Nesbit, J. C., Mahmoody Ghaidary, A., Jain, M., & Hajian, S. (2023a). Explanatory inferencing in simulation-based discovery learning: sequence analysis using the edit distance median string. *Instructional Science*, *51*(2), 309–

341. https://doi.org/10.1007/s11251-022-09614-4

Obaid, T., Aghajani, H., & Linn, M. C. (2023b). Using optimized clustering to identify students' science learning paths to knowledge integration. STEM Education Review, 1.

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology, 8*, Article 422. https://doi.org/10.3389/fpsyg.2017.00422

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*(1), 33–40. https://doi.org/10.1037/0022-0663.82.1.33

Platt, J. R. (1964). Strong inference. Science, 146, 347-353. Polya, G. (1954). *Mathematics and plausible reasoning: Vol. 2. Patterns of plausible reasoning*. Princeton, NJ: Princeton University Press.

Polya, G. (1954). *Mathematics and plausible reasoning: Vol. 2. Pat-terns of plausible reasoning*. Princeton, NJ: Princeton UniversityPress.

Popper. K. R. (1959). *The logic of scientific discovery*. New York:Basic Books

Quellmalz, E. S., M. J. Timms, M. D. Silberglitt, and B. C. Buckley. (2012). Science Assessments for All: Integrating Science Simulations into Balanced State Science Assessment Systems. *Journal of Research in Science Teaching 49*(3), 363–393. doi:10.1002/tea.21005.

Quintana, C., Reiser, B. J., Davis, E. A., Krajcik, J., Fretz, E., Duncan, R., … Soloway, E. (2004). A scaffolding design framework for software to support science inquiry. *Journal of the Learning Sciences, 13*(3), 337–386. https://doi.org/10.12 07/s15327809jls1303_4.

Reid, D. J., Zhang, J., & Chen, Q. (2003). Supporting scientific discovery learning in a simulation environment. *Journal of Computer Assisted Learning*, *19*(1), 9-20.

Reigeluth, C. M., & Schwartz, E. (1989). An instructional theory for the design of computer-based simulations. *Journal of Computer-Based Instruction, 16*(1), 1-10.

Reimann, P. (1991). Detecting functional relations in a computerized discovery environment. *Learning and Instruction*, *1*(1), 45–65. https://doi.org/10.1016/0959-4752(91)90018-4

Rieber, L. P., Tzeng, S.-C., & Tribble, K. (2004). Discovery learning, representation, and explanation within a computer-based simulation: finding the right mix. *Learning and Instruction*, *14*(3), 307–323. https://doi.org/10.1016/j.learninstruc.2004.06.008

Sanders, I., Bouwmeester, M., & Blanken, M. v. (2000). *Heuristieken voor experimenteren in ontdekkend leeromgevingen. [heuristics for experimenting in discovery learning environments]* (Unpublished report). Enschede: University of Twente.

Schauble, L. (1990). Belief revision in children: The role of prior knowledge and strategies for generating evidence. *Journal of Experimental Child Psychology, 49*(1), 31–57. https://doi.org/10.1016/0022-0965(90)90048-D

Schauble, L, Glaser, R., Raghavan, K., and Reiner, M. (1991). Causal models and experimentation strategies in scientific reasoning. *Journal of the Learning Sciences, 1*, 201-238.

Schnotz, W., Böckheler, J., & Grzondziel, H. (1999). Individual and co-operative learning with interactive animated pictures. *European Journal of Psychology of Education*, *14*(2), 245–265. https://doi.org/10.1007/BF03172968

Schraw, G., Crippen, K. J., & Hartley, K. (2006). Promoting Self-Regulation in Science Education: Metacognition as Part of a Broader Perspective on Learning. *Research in Science Education, 36*, 111-139. https://doi.org/10.1007/s11165-005-3917-8

Shah, A. K., & Oppenheimer, D. M. (2008). Heuristics made easy: An effort-reduction framework. *Psychological Bulletin, 134*(2), 207–222. https://doi.org/10.1037/0033-2909.134.2.207

Shute, V. J., & Glaser, R. (1990). A large-scale evaluation of an intelligent discovery world: Smithtown. *Interactive Learning Environments, 1*, 51-77.

Strand-Cary, M., & Klahr, D. (2008). Developing elementary science skills: Instructional effectiveness and path independence. *Cognitive Development*, *23*(4), 488–511. https://doi.org/10.1016/j.cogdev.2008.09.005

Swaak, J., van Joolingen, W. R., & de Jong, T. (1998). Supporting simulation-based learning; the effects of model progression and assignments on definitional and intuitive knowledge. *Learning and Instruction*, *8*(3), 235–252. https://doi.org/10.1016/S0959-4752(98)00018-8

Swann, W. B., Jr. (1984). Quest for accuracy in person perception: A matter of pragmatics. *Psychological Review, 91*, 451-All.

Sweller, J., Van Merrienboer, J. J., & Paas, F. G. (1998). Cognitive architecture and instructional design. *Educational psychology review, 10*, 251-296.

Tabak, I. (2004). Synergy: A complement to emerging patterns in distributed scaffolding. *Journal of the Learning Sciences, 13*(3), 305–335. https://doi.org/10.1207/s15327809jls1303_3.

Tsirgi, J. E. (1980). Sensible reasoning: A hypothesis about hypotheses. *Child Development, 51*, 1-10.

van Joolingen, W. R. (1999). Cognitive tools for discovery learning. *International Journal of Artificial Intelligence in Education, 10*, 385-397.

Van Joolingen, W. R., & De Jong, T. (1991). Supporting hypothesis generation by learners exploring an interactive computer simulation. *Instructional Science*, *20*(5/6), 389–404. https://doi.org/10.1007/BF00116355

van Joolingen, W. R., & de Jong, T. (1997). An extended dual search space model of learning with computer simulations. *Instructional Science, 25*, 307-346.

Veenman, M. V. J., F. J. Prins, and J. J. Elshout. (2002). Initial Inductive Learning in a Complex Computer Simulated Environment: The Role of Metacognitive Skills and Intellectual Ability. *Computers in Human Behavior 18* (3), 327–341. doi:10.1016/S0747-5632(01)00038-3.

Veermans, K. H., de Jong, T., & van Joolingen, W. R. (2000). Promoting self directed learning in simulation based discovery learning environments through intelligent support. *Interactive Learning Environments, 8*, 229-255.

Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology, 12*, 129-140.

Wason, P. C, & Johnson-Laird, P. N. (1972). *Psychology of reasoning*. Structure and content. London: Batsford.

Wiese, & Linn, M. C. (2021). It Must Include Rules. *ACM Transactions on Computer-Human Interaction, 28*(2), 1–41. https://doi.org/10.1145/3415582

Winne, P. H. (2022). Modeling self-regulated learning as learners doing learning science: How trace data and learning analytics help develop skills for self-regulated learning. *Metacognition and Learning*, *17*(3), 773–791. https://doi.org/10.1007/s11409-022-09305-y

White, B. Y. (1993). ThinkerTools: Causal models, conceptual change, and science education.*Cognition and instruction, 10*(1), 1-100.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry, 17*(2), 89–100. https://doi.org/1 0.1111/j.1469-7610.1976.tb00381.x.

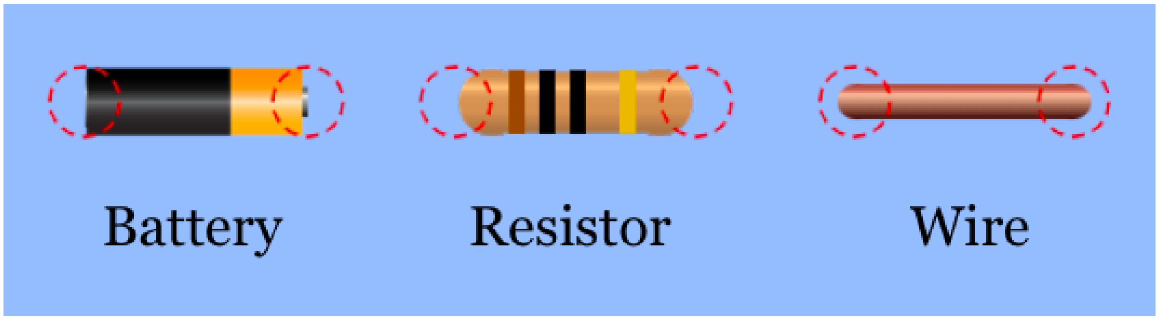
Zimmerman, C. (2000). The development of scientific reasoning skills. Developmental Review, 29, 99-149.

Zohar, A. (1995). Reasoning about interactions between variables. *Journal of Research in Science Teaching, 32*, 1039-1063.

Appendix A. Pretest and Posttest

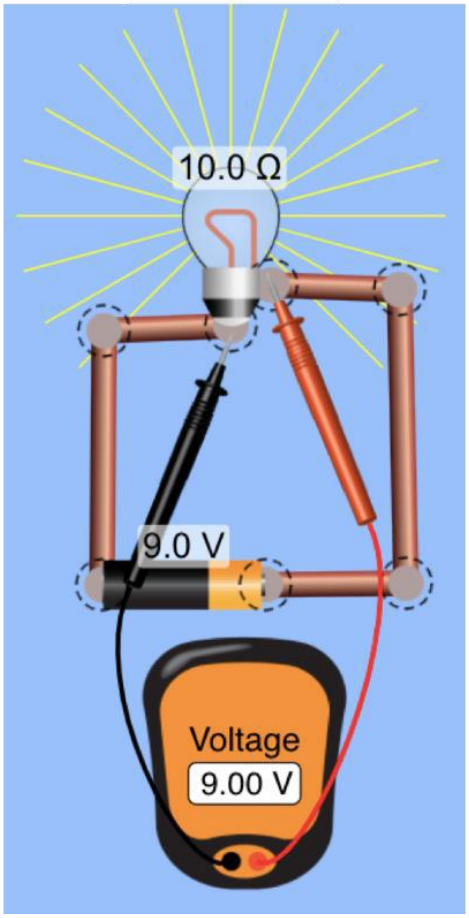
Please answer the following questions.

These elements are used in each question:

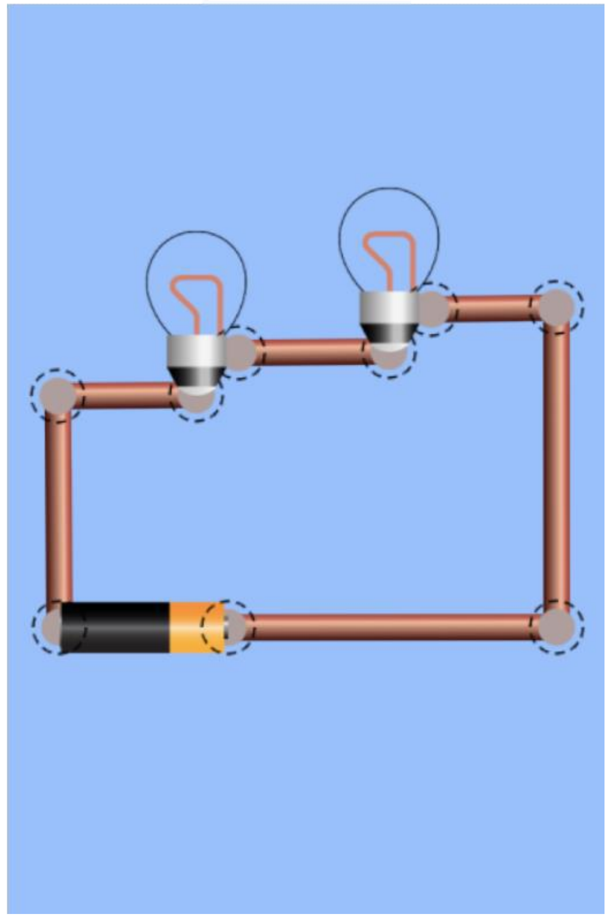


1. Cannice created Circuit A and tested the voltage drop. Amy asked her what would happen to the voltage drops in Circuit B. Based on her evidence from Circuit A, Cannice concluded: "In a series circuit, the voltage drop will be the same across each individual resistor/light bulb."

Circuit A



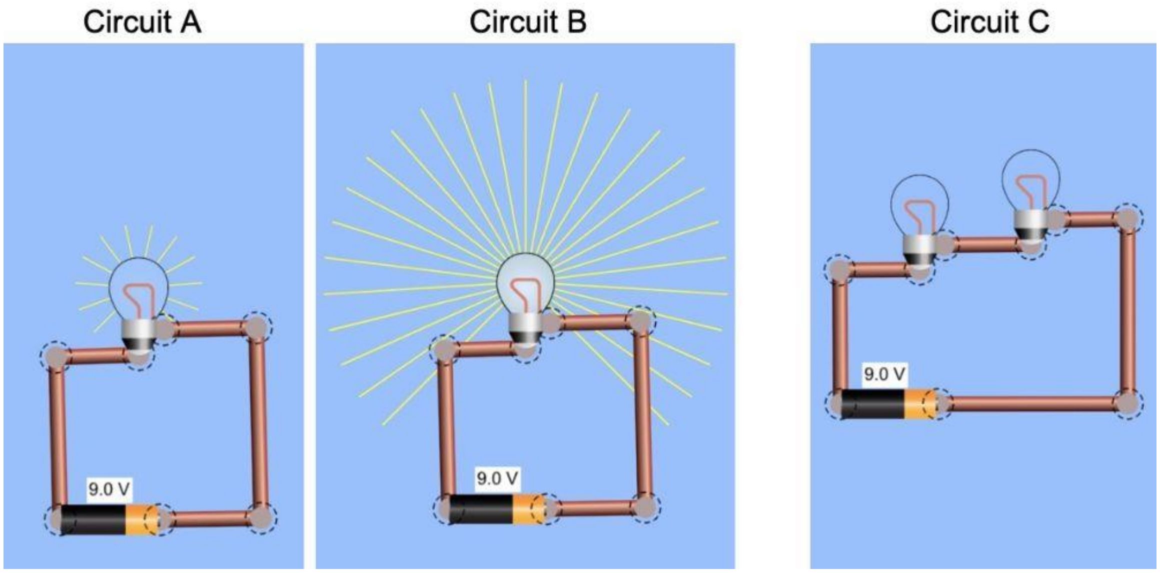
Circuit B



1.1. Select the best answer.
A. Cannice's conclusion is correct.
B. Cannice's conclusion is incomplete.
C. Cannice's conclusion is incorrect.

1.2. Explain why you selected A, B, or C?

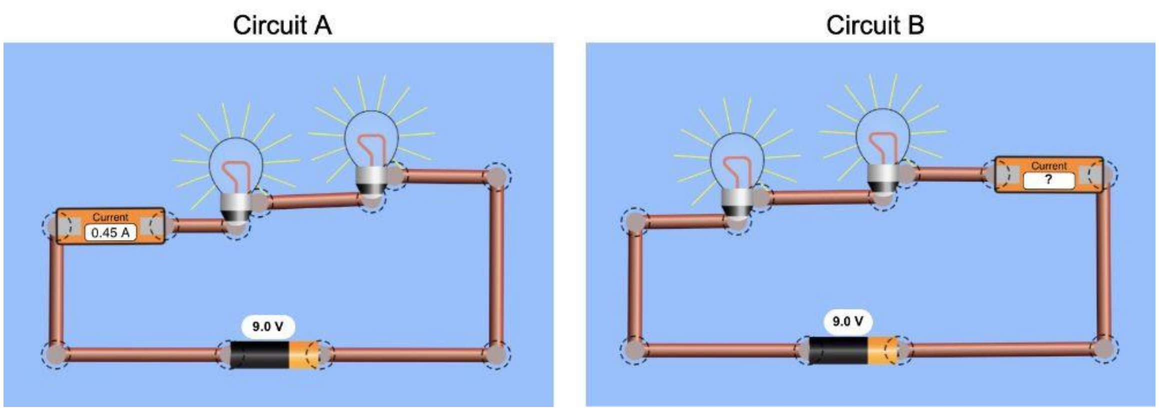
2. Amy created the series circuits below. Based on the evidence from Circuit A and Circuit B, she concluded about Circuit C: "In a series circuit, the current in each individual resistor/light bulb will always be different."



2.1. Select the best answer.
A. Amy's conclusion is correct.
B. Amy's conclusion is incomplete.
C. Amy's conclusion is incorrect.

2.2. Explain why you selected A, B, or C?

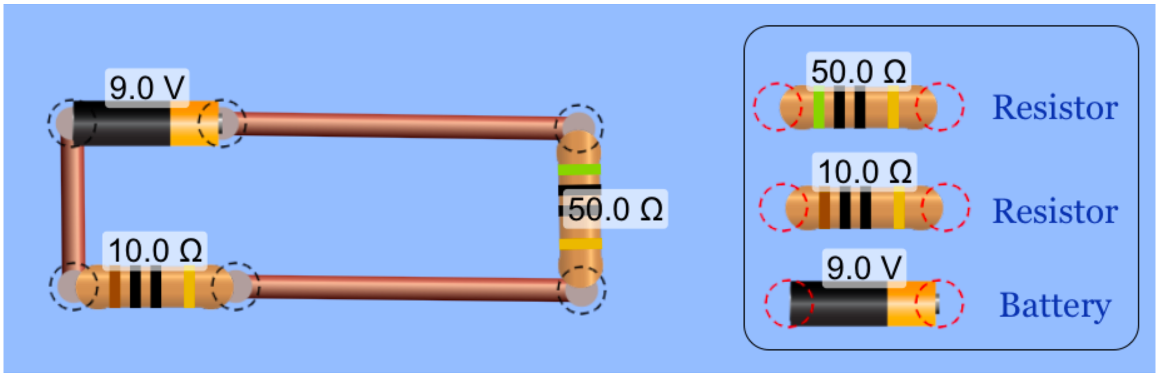
3. Carmine created the series circuits below. Based on the evidence from Circuit A, he concluded about Circuit B: "the current will decrease." (Note: current is moving clockwise).



3.1. Select the best answer.
A. Carmine's conclusion is correct.
B. Carmine's conclusion is incomplete.
C. Carmine's conclusion is incorrect.

3.2. Explain why you selected A, B, or C.

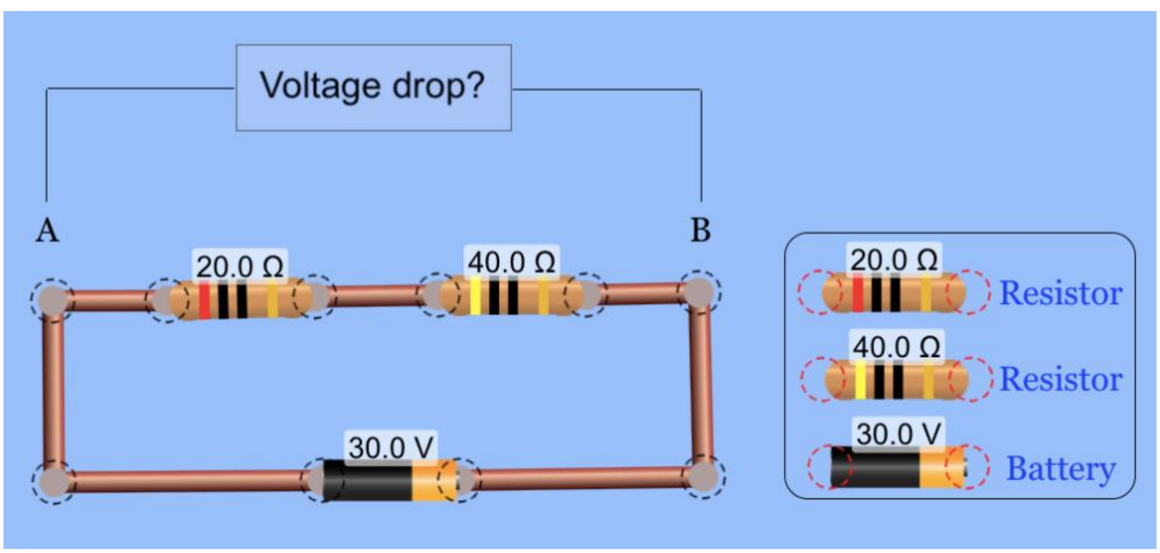
4. A battery of 9 volts (V) powers a series circuit with two resistors. One resistor has the value of 50 Ohms (Ω) and the other resistor has the value of 10 Ohms (Ω). What will happen to the current in each resistor?



4.1. Select the best answer.
A. The current will be reduced more in the 10 Ohms resistor. And, the current will be reduced less in the 50 Ohms resistor.
B. The current will be reduced less in the 10 Ohms resistor. And, the current will be reduced more in the 50 Ohms resistor.
C. The current will be the same in both resistors.

4.2. Explain why you selected A, B, or C.

5. In a series circuit, we have a battery of 30 volts (V) and two resistors. One resistor has the value of 20-ohm (Ω) and the other resistor has the value of 40-ohm (Ω). What will happen to the voltage drop between points A and B?

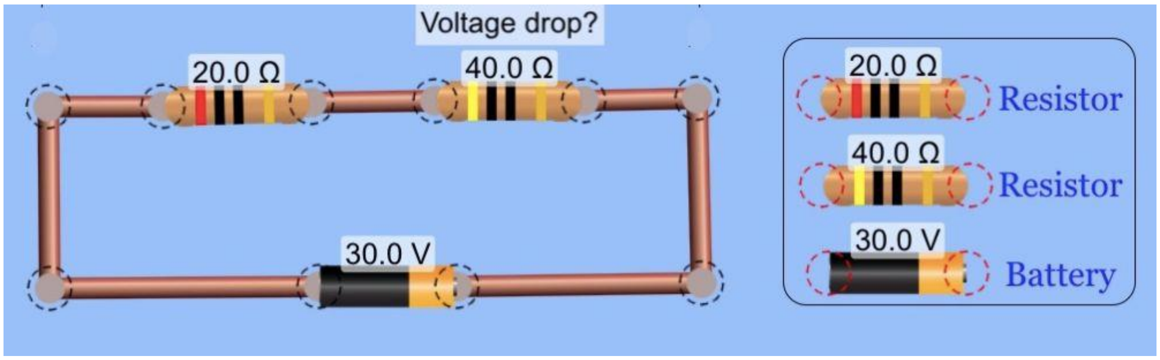


5.1. Select the best answer.
A. Between point A and B, the voltage will be lower than 30 volts.

B. Between point A and B, the voltage will be 30 volts.
C. Between point A and B, the voltage will be higher than 30 volts.

5.2. Explain why you selected A, B, or C.

6. In a series circuit, we have a battery of 30 volts (V) and two resistors. One resistor has the value of 20-ohm (Ω) and the other resistor has the value of 40-ohm (Ω). What will happen to the voltage drop in the resistor with 40 ohms (Ω)?



6.1. Select the best answer.
A. Voltage drop in 40-ohm (Ω) resistor will be higher than the voltage drop in 20-ohm (Ω) resistor.
B. Voltage drop in 40-ohm (Ω) resistor will be lower than the voltage drop in 20-ohm (Ω) resistor.
C. Voltage drop in 40-ohm (Ω) resistor will be the same as the voltage drop in 20-ohm (Ω) resistor.

6.2. Explain why you selected A, B, or C.