Title
Designing Process-Oriented Computational Assistance to Support Self-Regulated Learning in Complex Games

Author
Kleinman, Erica Michelle

Publication Date
2023

UNIVERSITY OF CALIFORNIA
SANTA CRUZ

DESIGNING PROCESS-ORIENTED COMPUTATIONAL ASSISTANCE TO SUPPORT SELF-REGULATED LEARNING IN COMPLEX GAMES

A dissertation submitted in partial satisfaction of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

in

COMPUTATIONAL MEDIA

by

Erica M. Kleinman

June 2023

The Dissertation of Erica M. Kleinman
is approved:

Dr. Magy Seif El-Nasr, Chair

Dr. Edward Melcer

Dr. Casper Harteveld

Dr. Norman Makoto Su

Peter Biehl
Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures

List of Tables

Abstract

Designing Process-Oriented Computational Assistance to Support Self-Regulated Learning in Complex Games

by

Erica M. Kleinman

Complex games, those with multiple correct strategies and unpredictable outcomes, are seeing increased popularity and integration into high-impact domains such as health, education, and training. Further, even in entertainment contexts, these games have proven benefits for players. The steep learning curves, however, make the games inaccessible to many players, resulting in a lack of diversity in professional circles, cutting people off from the proven benefits of play, and rendering them ineffective in serious domains.

This work examines, in a user-centric manner and through the lens of the Cyclical Phase Model of Self-Regulated Learning, how players learn to play and master complex games and how we can better support these processes through computational support. These thesis, thus, contributes an advanced understanding of how learning and mastery occurs in complex gameplay, and empirical insights into how to support these processes, grounded in theories of learning.

In sum, this work contributes to making complex games more accessible and effective as high-impact tools and identifies generalizable insights for the design of computational support tools for learning relevant to adjacent domains including explainable and user experience of Artificial Intelligence.

To my Grandmother, Anita.

She earned a PhD, despite unspeakable adversity.

Her memory became my motivation.

Acknowledgments

While I alone write this thesis, I would not have made it here, both literally and figuratively, were it not for the help of countless others. Here, I take the opportunity to thank those who have contributed to my success and the creation of this document.

Magy Seif El-Nasr, my advisor, for providing me with valuable guidance, feedback, support, and motivation throughout my entire time as her student. For letting me explore the topics I was interested in, but also for preventing me from going down the wrong path. And of course, for taking a chance on me and giving me the greatest opportunity of my life.

My committee, Edward Melcer, Casper Harteveld, and Norman Makoto Su, for their continued support, guidance, and insight, and for agreeing to read not only this document, but my proposal, despite how long it was.

All of the members of the GUII lab, both on the East and West coasts, for their continued friendship and collaboration and for creating a welcoming and supportive community.

Elin Carstensdottir, for being the greatest senpai I could have asked for, a wonderful collaborator and co-author, and a good friend. And for letting me show her lots of terrible anime.

My friends on the East Coast, for their long-distance support, and especially to my roommates, Mel, Hannah, and Greg, for letting me remain a part of their household, even after I moved to the other side of the country.

My friends on the West Coast for welcoming me into their circles and giving me a place to belong and be myself when I was so far from everything I knew. And especially to Ruby, Victor, and Stephen for many wonderful card games.

My family, for always supporting me and letting me pursue my dreams and my interests, even when they did not fully understand them, and for calling me every night to make sure I was ok. Special shout out to my brother, Matthew, for driving across the entire country with me when I moved to California.

My classmates at Northeastern and UCSC, for creating a fun and supportive environment.

The students I mentored and taught during my time as a PhD student, for giving me confidence in my ability to teach.

And finally, I would like to thank all the horses I have gotten to ride during my PhD, though they will never read this, their presence kept me sane.

Part I

Introduction and Background

Chapter 1

Introduction

1.1 Research Motivation

Complex games, those that have multiple correct strategies or solutions and/or unpredictable situations or results, are seeing greater use in serious and high-impact domains including health [172], education [321, 382], military [155], and disaster response [602]. This is because the games themselves are engaging to users, making them a more interesting and, in many cases, a more hands-on, alternative to traditional education or training methods. Further, high-skill level play in complex games has demonstrated proven benefits for the players, even in entertainment focused games. The mechanics of dynamic, fast-paced complex games have demonstrated an ability to nurture fine motor skills among players [673, 520, 289] and improve their reaction time [360]. Additionally, while not all complex games are multiplayer, those that are have demonstrated an ability to improve players' social-emotional skills, including communication, teamwork, and emotional regulation [727, 362, 479, 675]. Complex gameplay, especially in esport games, has also demonstrated the ability to improve younger players' academic performance, by nurturing their critical thinking and problem solving skills [552, 120, 382]. Further, esports have grown into a multi-billion dollar industry with thousands of career opportunities for players capable of high-skill level play [15].

The problem, however, is that complex games also tend to be notoriously difficult to learn, rendering them almost inaccessible to large portions of the population who do not already belong to communities or have support networks that can help them improve. While the games themselves offer tutorials that teach players the basics of interaction, the skills necessary to succeed in complex gameplay situations, where there may be multiple pathways to success and unexpected outcomes, are rarely taught or coached by the games themselves. As a result, players must learn and improve at play on their own, but the path to success is not always clear. It can be easy for players to make the wrong choices, fail to understand what they are doing wrong, and ultimately become frustrated and quit playing when they fail to meet their own expectations [88, 196].

This means that the benefits of play, such as fine-motor or communication skill improvement, may be inaccessible to many players, simply because they are unable to reach the necessary level of play on their own and do not have a support network that can help them. Further, those games that are intended for use in serious domains may be rendered ineffective if the players meant to benefit from them are unable to even play them. In situations like the esports industry, this can also result in a lack of diversity, as those players most able to succeed are those who belong to a community or have a support network that can help them learn through example, coaching, and advice. Players from different locations or of different demographics from those represented by existing support networks, however, face a far greater challenge in the pursuit of elite play [660] and may never be able to reach the highest skill levels as there is little support available to help them do so.

In order to make high-skill level play in complex games more accessible to more players, we need to better facilitate players' ability to learn and improve on their own. But in order to do that, we require a stronger understanding of how players learn and improve at complex gameplay on their own. Within the learning sciences, Self-Regulated Learning (SRL) is a learning theory that specifically formalizes the cognitive, meta-cognitive, behavioral, motivational, emotional, and affective aspects of learning on one's own [497]. In other words, SRL broadly refers to the phenomenon by which students can self-regulate their learning process without the direct guidance of an educator [497, 757].

Learning gameplay, specifically learning the skills necessary to move from the basics of interaction to high-level play, is almost exclusively learned in the absence of a formal educator, with players dependent on their own goals and reflections [480, 88]. Thus, SRL, as a learning theory, lends itself well to understanding and formalizing learning in complex games.

Further, I focus here on Zimmerman's Cyclical Phase Model of Self-regulated Learning (CPM) [750]. CPM splits the processes of SRL into three phases: forethought, performance, and self-reflection [497, 757, 756]. Forethought encompasses skills used to plan or set goals for a learning activity, performance encompasses skills used to complete the activity and monitor one's progress towards goals, and self-reflection encompasses skills related to evaluating one's performance and adapting it for future iterations of the activity. As the model is a cycle, executions of self-reflection inform future executions of forethought. This model lends itself particularly well to complex games, where players are often required to undergo the same task in repeated cycles, with opportunities to plan, monitor, and self-reflect.

For example, role playing games will require players to battle powerful monsters. Players will go into these fights with a plan, and bring certain weapons, characters, or items to enact that plan. During the fight, players will monitor their performance and the execution of their chosen plan. After the fight, players will reflect on what worked and what did not, and adjust their plans, either because they need to try again, or because they will have to fight another powerful monster later. Esports present an even more direct application of CPM, with clearly defined pre-game, in-game, and post-game time-points that correspond to forethought, performance, and self-reflection.

Previous work has explored SRL in games, however, it has focused almost exclusively on educational games and the impact of SRL on the extent to which players learn the educational material behind the game [480, 481, 571]. Work that explores the impact of SRL skills on learning gameplay itself is, in contrast, quite limited [480, 88]. As a result, there is currently a very restricted understanding of how SRL manifests in the context of learning complex gameplay. This, in turn, means that supporting SRL, and specifically SRL in terms of CPM, in complex gameplay remains an open question.

The problem I seek to address in my dissertation then is how to best support SRL in terms of CPM for players who do not necessarily have a group they may turn to for aid. In the absence of other humans, data, which has become increasingly available in recent years with advances in technology [185], has proven itself a promising avenue for supporting SRL in complex games. Many games have already integrated visualizations and summaries of gameplay data to help players learn through review, reflection, and comparison with others [280, 80, 449, 448, 447]. Further, outside of the games themselves, many computational tools have been developed with the explicit goal of allowing players to view, interact with, and interpret their data [701, 11, 367, 14, 706].

That being said, to date, there has been little empirical examination of how data and data-driven tools can be used to enhance learning, with those studies that do exist taking a game or context specific approach [708, 509]. Instead, much of the empirical study of player-facing data focuses on design and usability [701, 274, 703, 706]. Further, while supporting SRL through computational assistance is common in learning sciences [497], there is almost no literature on the topic in the context of games, especially in the context of learning the gameplay itself.

In this dissertation, I argue that we can support players in learning complex gameplay, and ultimately improve their experience, by using gameplay data to enhance the execution of self-regulated learning skills in the context of play. In doing so, we can make the benefits of play accessible to more players, which will, in turn, increase diversity in complex gaming domains and make complex games in high-impact areas more usable and effective. I explore this space through user-centric studies of how players learn, the execution of SRL within the gameplay learning process, how it is supported by the current state of the art of data-driven tools, and how we can advance the design of these tools to enhance that support. In the following subsection, I will summarize the outcomes of my dissertation work more specifically.

1.2 Research Questions

The overarching research question I seek to answer is as follows:

- **Overarching Research Question (ORQ)**: How can we facilitate, enhance, and encourage self-regulated learning in the context of improving at complex gameplay?

In answering this research question, I recognize four research thrusts in this dissertation. An overview of these thrusts can be seen in Figure 1.1. The first two research thrusts look broadly at SRL in games as defined by CPM. Through mixed-methods techniques, including interviews and surveys, I explored how players learn and improve at play in the context of CPM and how gameplay data can support that process. These are followed by two thrusts that explore the impact of data, and specifically process visualizations, on SRL and CPM in complex games, again, through various user studies. I will briefly introduce these thrusts and their related research questions in the following subsections.

1.2.0.1 Thrust 1: Studies of Self-Regulated Learning in Complex Games

The first thrust examines learning and SRL in the context of complex games, specifically esports, and sought to answer the following research question:

- RQ 1: How do players engage Self-Regulated Learning skills in the context of learning and improving at play?

I explored this question through three studies. The first, an interview study, asked esports players about their goals, practice routines, and experiences trying to improve at their respective games. It followed this with a hypothetical design exercise asking players to design a fictional computational tool that could help them improve at their game. The results revealed four activities and four challenges that players face when trying to improve.

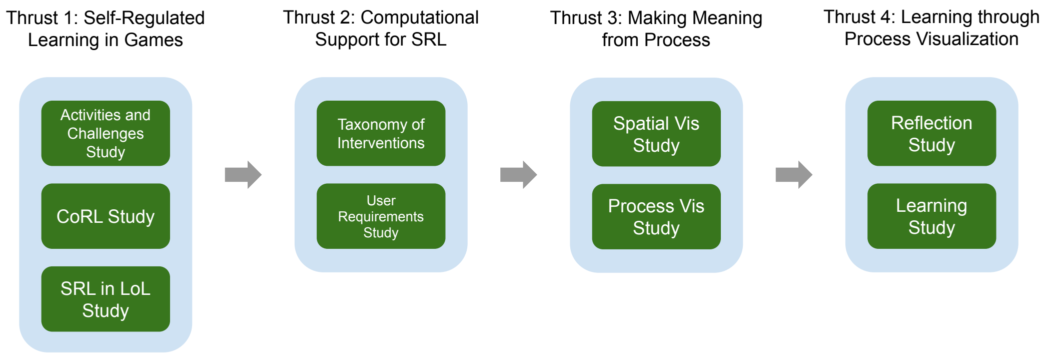


Figure 1.1: An overview of the four research thrusts of this dissertation.

The second study briefly explored the question of social learning in esports play and leveraged Hadwin et al.'s theory of Co-Regulated Learning [273] to understand how esports teammates co-regulated each other's learning processes. This was another semi-structured interview study, the results of which revealed 18 themes regarding how players regulate each others' learning process and considerations for the design of tools that may aid in this co-regulation process by supporting players the way other players do. While this dissertation work primarily focuses on solo learning, this study sets the first stones for future work that explores the question of social learning more thoroughly.

The third study replicated the work of Kitsantas and Zimmerman [343] and looked at how SRL skills were executed by *League of Legends* players at various skill levels. The results revealed statistically significant differences in the forethought phase, but not in the performance or self-reflection phases, suggesting that players across skill levels were executing SRL skills evenly in these phases. Based on these results, I concluded that the data visualizations available to players during these phases encourage SRL skill execution, regardless of skill level.

In sum, the primary takeaway of thrust 1 was an empirical understanding of how learning occurs in complex games, both generally, and specifically in terms of CPM.

1.2.0.2 Thrust 2: Supporting Self-Regulated Learning Skills in Complex Tools through Computational Support

Based on the findings of the first thrust, I concluded that gameplay data, often made apparent through interface elements, provides players of complex games, especially esports, with some manner of support for SRL skill execution. Thus, the second thrust of this work sought to answer the following research question:

- RQ 2: How do data-driven tools support self-regulated learning skills in complex games?

To answer this question, first, I worked with collaborators to conduct a systematic review of existing support tools for esport games. This review revealed nine intervention types offered by the tools and patterns of when during gameplay they are offered, mapped to the three phases of the Cyclical Phase Model. I present this as a taxonomy of intervention types for computational assistants for esports.

Following the creation of the taxonomy, I conducted a two-part, mixed methods survey and interview study that gauged players' preferences in terms of when during the gameplay experience they would want to interact with each intervention type. Using an explanatory sequential method, preferences were first collected using a survey. These survey results were then expanded on through qualitative interviews with a subset of respondents. The results revealed usage patterns and preferences from players regarding how existing features support SRL skills across the three phases for complex games and opportunities for better support.

In sum, the primary takeaways of thrust 2 are a taxonomy of the types of interventions offered by computational support tools for esports and an overview of how existing computational tools support SRL in terms of CPM, and opportunities to improve that support based on user input collected through a mixed-methods study.

1.2.0.3 Thrust 3: Making Sense of Visualizations of Process

Based on the findings of the previous two thrusts, I developed a solid understanding of how players of complex games leverage SRL skills towards improvement, and how they use data towards enhancing this process. From these results, I recognize two facts. First, reviewing gameplay, and gameplay data, post-play, during the self-reflection stage, is one of the most common executions of SRL in the context of improving at a complex game. The self-reflection phase encompasses meta-cognitive processes related to evaluating performance, attributing failure, and revising plans and goals for the future based on reviews of one's actions after the fact [756, 75, 273].

In the learning literature, reflection has been shown to be critical to effective self-regulated learning, and many methods and tools have been developed to prompt and encourage student reflections [497]. Reflection has also been demonstrated as a key element of gameplay experiences, where it can help players come to new conclusions about the narrative or their own actions [453]. In learning games, previous work demonstrated how prompting reflection can improve the rate at which players learn educational material [693, 569]. The results of the previous two thrusts further suggest that reflection also plays a key role in players learning and improving at the gameplay itself.

Second, being able to understand the causal relationships between in-game actions is recognized, by many players, as critical to successful improvement over time. However, existing data visualization interfaces for self-reflection do not sufficiently provide enough information to understand causality, as they are primarily presenting aggregate statistics. This makes it difficult for players to understand where their own gameplay went wrong, and also makes it difficult to learn from the gameplay of others [625, 112], as it is almost impossible to glean strategic information from aggregate data.

Based on this understanding, process visualizations, those that visualize data in a granular, action by action manner, may enhance self-regulated learning during the self-reflection phase by providing more causal information. Process visualizations are common in other domains that focus on formalizing human behavior, such as process mining [684, 551, 718]. These visualizations all follow a similar approach, using node-link diagrams to display the ordering and progression of human actions as they work towards completion of a task [514, 718, 687, 581]. This visualization approach makes them better suited for extracting causal relationships from data, as the progression of one event to another is clearly depicted. While similar visualizations exist within the domain of games, they have been used almost exclusively for game user research rather than as player-facing tools [478, 702, 26]. This is likely due to the, naturally, more complicated nature of process visualizations, which may render them difficult for players to make meaning from.

Thus, the third research thrust sought to explore the extent to which players can extract meaningful insights from gameplay data presented in a process visualization. Thus, this thrust explores the research question:

- RQ 3: How do players of complex games extract meaningful insights from visualizations of process?

This question was explored through two studies. The first was a qualitative study that examined how players of the esport game *DotA 2* made meaning from spatio-temporal visualizations of others' gameplay data. The results revealed a preliminary interaction taxonomy for the domain as well as a process model.

The second study built on the results of the first, but looked instead at a sequential-process visualization. The results revealed two methods for sense making using sequential-process visualizations in the context of complex games.

In sum, the primary takeaways of thrust 3 are initial interaction taxonomies and process models for how players make sense of game data visualized in a process sensitive manner.

1.2.0.4 Thrust 4: Learning through Process Visualizations

Thrust 3 demonstrated that players are able to extract meaningful insights from process visualizations and that comparison with others' data influenced one's perspective and interpretation of their own data and experience. The final question I sought to explore in Thrust 4 was thus:

- RQ 4: How do process visualizations of one's own and others' gameplay data impact self-reflection and learning?

To answer this question, I first examined the impact of others' process data on one's reflection on their own data through a mixed methods study. The results revealed that comparison with peers significantly influenced a given player's willingness to try a different approach if they repeated the task, without having any negative impact on the quality of reflection.

I then conducted a second study examining the impact of process visualization on SRL processes in the self-reflection phase of CPM, when compared to an aggregate visualization. The results revealed that players who reflected on a process visualization had a significant improvement in performance and were significantly more likely to evaluate their performance based on their use of the correct method or strategy rather than just their score.

In sum, the primary takeaways of thrust 4 are the empirically derived relationships between process visualizations of one's own and others' data and self-reflection and performance. From these, I derive suggestions for the design of computational assistants that I discuss further below.

1.3 Contribution Summary

To briefly summarize, the goal of this dissertation is to develop an understanding of how players learn in complex games and guidelines for how we can design computational assistants to better support that learning through the theoretical lens of the Cyclical Phase Model of Self Regulated Learning. I pursue this goal for several reasons. First, complex games have proven benefits for players, which become inaccessible to those who are unable to learn and master gameplay. Second, some complex game communities, such as esports, have developed a diversity issue due to this inaccessibility. Third, complex games are seeing increased use in serious domains, where it is critically important to ensure that players are able to learn and master gameplay in order for the game to achieve its design goals.

Through this dissertation, I have accomplished the following towards the goals I discuss above, all of which have been published or are currently under review at major publication venues:

- Identified common activities and challenges experienced by esports players when they are attempting to learn to play
- Highlighted similarities and significant differences in how skills from the cyclical phase model are leveraged across expertise levels in a complex gaming context and identified the connection between these skills and computational support elements
- Developed a taxonomy of interventions offered by computational assistants for esports games
- Identified how esport players use computational assistants to support SRL and opportunities to improve this
- Highlighted players' reservations towards computational assistance in a complex gaming context and opportunities to address this through design
- Developed two taxonomies, as well as process models, for interaction with gameplay data, one for spatio-temporal data and one for process visualizations
- Determined that comparison with peers' process data can increase a player's desire to try a different approach in future play while not reducing the quality of their reflections

- Determined that reflecting on one's own process data, as opposed to aggregate data, has a significant impact on evaluation and performance.

Chapter 2

Background and Related Work

2.1 Complex Games

Games exist across a wide spectrum in terms of aesthetics, platforms, and gameplay. Some games are more casual, restricting themselves to a small number of mechanics, a single way to complete an objective, and predictable outcomes. Other games, however, are what people would refer to as "complex". But what does it mean to say a game is complex? In a vacuum, complexity refers to the presence of many different and connected parts. In the context of games, however, there exists no single definition for a "complex" game.

In some cases, the literature measures complexity by the size and number of objects or interactions present in a game [169, 27]. In other cases, the phrase "complex" is used with no explicit definition, and simply refers to games that are known to consist of many "moving parts" [659].

A more comprehensive distinction is provided by Prensky [536], who defines complex games as those that takes tens of hours to play, demand that the player learn and master multiple skills, and research and communicate outside the game. Prensky goes on to say that complex games often require players to make decisions and take action in real time and contain ethical dilemmas. Prensky ultimately emphasizes the role that learning and mastering skills, realized as one progresses through game levels as the key characteristic of the complex gameplay experience.

In the context of this dissertation, I use the term “complex games to refer to those games that have multiple correct strategies or solutions that a player may choose from and unpredictable outcomes or results. An example of a complex game that illustrates this is *The Wolf Among Us*, which features a scene in which the player character enters a bar to speak with the suspect of a crime. The player is able to choose a friendly approach to the conversation or an adversarial one, both of which are valid methods to progress the scene (multiple correct strategies or solutions). What the non-player character (NPC) says or does in response, however, may not reflect the route the character takes. For example, even if the player takes the friendly route, the scene can still end in a bar fight (unpredictable outcomes or results).

Another example is the mobile game *Disney Twisted-Wonderland* where players put together teams of five characters from a selection pool to fight enemies in turn-based combat. There is no one right team for any given fight and players can adopt any team combinations they wish based on what is available to them (multiple correct strategies or solutions). Enemies have various skills that may be strong or weak against the player's team's skills and on any given turn the player must choose at least one skill to use without knowing what the enemy will use, meaning the skill they choose could be very effective or very ineffective (unpredictable outcomes or results).

These factors, however, can make complex games initially difficult to master, as it is not always clear what the correct course of action is and, when something goes wrong, it can be difficult to pinpoint the cause. Nevertheless, complex games can be mastered through practice and continued exposure, as the player becomes familiar with the space of possible actions and outcomes and builds causal relationships between the two.

Complex games are not restricted to any one style or genre, and encompass anything from *Tetris* to *The Legend of Zelda, Breath of the Wild.* Notably, complex games also encompass multiplayer games, which also increase the complexity by adding interactions with another player, and can range from *Final Fantasy XIV* to *Chess.* In the former, the other players are typically on the same side and in the latter they are on the opposing side. There still remain multiple correct strategies and unpredictable outcomes at any given moment but there now exists a level of unpredictability in not knowing how the other player or players will behave. Many educational games, such as *May's Journey* [321], also count as complex games under this definition as they are often designed with multiple possible solutions, to encourage students to think and experiment for themselves.

That being said, complex games, especially entertainment ones, are, notably, often meant to be a challenge, as is the case for the *Dark Souls* series, for example. Players who choose to play these games are typically doing so to enjoy the challenge. This thesis work is not meant to remove the challenge from challenging games, but instead, is meant to help those players who wish to enjoy the game but find the challenge to be a road-block to their enjoyment. A seemingly insurmountable challenge, or as Esteves et al. [196] found, perception that one is performing poorly compared to others, can lead to discontinuation of play. This outcome is undesired, as it will keep players from experiencing the aforementioned benefits of play and, on the business side, may prevent them from buying other games from the franchise in the future.

While many have examined dynamic difficulty adjustment as a way to cater the challenge to the skills of the individual player [309, 762], not all complex games, especially online ones, may employ such a system. The goal of this work is, thus, to provide insights into how to develop assistants that may help those players who find the challenge of learning a game to be insurmountable on their own. Similar to the customizable auto-chips of *Nier: Automata*, which allow the game computer to automatically operate some elements of the combat kit, these assistants are here for those who wish to use them, and may be ignored by those who enjoy the challenge. Additionally, computational assistants are about helping the player learn to rise to the challenge of the game by examining and understanding their own behavior. In doing so, they help the player adapt to the challenge, rather than remove the challenge from play.

2.2 Theories of Learning

In his 2012 book "Learning Theories: An Educational Perspective" [590], Dale Schunk defines learning as "an enduring change in behavior, or in the capacity to behave in a given fashion, which results from practice or other forms of experience." (P. 5). He goes on to discuss how learning involves changing through experience and that it endures over time. He clarifies, however, that there is no single definition of what learning is. That being said, the definition he provides certainly suggests that players are learning when they are improving at gameplay. Similarly enigmatic are the theories of how people learn, of which there are many. In this section, I will discuss the four theories of human learning that Schunk details in his book [590].

An early theory of learning, according to Schunk [590], was conditioning theory, which understood learning, specifically learning behavior, in terms of environmental events. This theory argued that learning occurred through the repeated reinforcement of desirable responses to external stimuli [268] and was built on the idea that humans learn by building connections between experiences and responses, often through trial and error [450]. According to this theory, positive reinforcement, i.e. new or continued stimuli, will lead to prolonged behavioral change, i.e. learning, while extinction of a behavior can occur if reinforcement is not received [590]. For example, if students who raise their hands are called on they will continue to do so, but if they are not called on, they will stop raising their hands. In educational contexts, this theory informed learning environments with strict progressions from one task to another, and immediate feedback that would prompt the learner to repeat a task if done incorrectly the first time [590].

While common in the early half of the 20th Century, conditioning theories waned in popularity in the 1950's and 60's due to increased awareness of the social and cognitive aspects of learning [590]. This awareness led to the rise of social-cognitive theory, which argued that people could acquire knowledge, rules, skills, strategies, beliefs, and attitudes by simply observing others, and not exclusively by responding to stimuli themselves [590]. What specifically sets social cognitive theory apart from its predecessor is that learning occurs both by performing tasks and by observing others perform tasks [592, 42]. In addition, according to this theory, through interaction with others and their environments, humans learn about the appropriateness of actions and their relative skills, which build cognitive beliefs about their capabilities and expected outcomes, which, in turn, inform future action [592, 42]. These beliefs, an element of self-efficacy, are considered an important part of education, and learning environments that subscribe to this theory will work to build students' self-efficacy to ensure that they will take the initiative to perform learning tasks [590]. Unlike conditioning theory, which has seen little discussion in recent literature, social cognitive theory is still relevant today, with recent work referencing it and combining it with other theories of learning [709, 485, 9].

Another learning theory that challenged the assertions of conditioning and behaviorism is information processing theory, which asserts that humans encode new information, relate it to known information, store it in memory, and retrieve it as needed [590, 614]. The general idea of this theory, which is actually a collection of theories, is that information is processed by the learner between receiving a stimulus and producing a response [590], which is in contrast to conditioning theory, in which it was argued that humans just responded to stimuli and learned through conditioned reinforcement. In educational contexts, information processing theories have prompted educators to connect newly taught information to prior knowledge and articulate explicit connections between different pieces of knowledge. These theories also relate to theories of attention [277] and cognitive load [648], in that they formalize the processes by which human learners take note of, store, and manage incoming information.

There is also the theory of constructivism, which shifts the focus from how knowledge is acquired to how it is constructed by learners [590]. Constructivism asserts that knowledge is built internally, rather than obtained from an external source, and therefore subjective and unique to individual people [590]. In educational contexts, constructivist theories have prompted a shift towards student-focused learning, where students are actively involved in guiding the learning process, rather than passive listeners [36]. Constructivism also posits that learners are not blank slates and instead bring pre-existing knowledge with them into a learning environment [474]. A number of contemporary learning techniques facilitate constructivist learning including discovery learning, where students obtain their own knowledge through problem solving, inquiry teaching, in which teachers question students, and discussion and debate-based teaching, which allow students to share multiple viewpoints on a topic [590].

The above are the four foundational theories of how humans learn as discussed by Schunk [590]. There exist various teaching and learning techniques which build upon these theories, such as experiential learning, which exists at the intersection of behavioral and cognitive theories, claiming that learners learn through a combination of concrete experiences and abstract conceptualizations of experiences [359]. One of the most influential techniques, however, is Self-Regulated Learning [590, 750], which details the processes by which a learner can self-regulate their use of the various experiential, behavioral, cognitive, and meta-cognitive processes involved in learning in order to drive their own learning process. What differentiates Self-Regulated Learning is, among other details, its lack of reliance on an external educator. In the following section, I will discuss Self-Regulated Learning in more detail.

2.3 Self-Regulated Learning and The Cyclical Phase Model

2.3.1 Self-Regulated Learning

Self-Regulated Learning (SRL) is the active and goal-directed process in which learners are portrayed as active and intentional regulators of their own cognition, metacognition, motivation, and behavior [416]. In other words, SRL refers to the phenomenon by which learners can learn on their own, in the absence of an educator. As such, SRL encompasses the processes and skills related to analyzing tasks, setting goals, developing strategies to reach those goals, monitoring progress towards those goals, and reviewing performance and outcomes [497, 540, 369].

I choose to focus on SRL, as a learning theory, in this work, due to its relevance to the phenomenon of learning gameplay. Specifically, gameplay, especially in entertainment games, is typically learned in the absence of a formal educator, such as a coach. Players are, thus, reliant on their own ability to to set goals, monitor progress, and review performance in order to identify failures and opportunities for learning and improvement. In other words, players, by nature, are required to self-regulate their learning when learning complex gameplay, especially when doing so in the absence of a formal educational setting.

There are several different theories and models of SRL [497, 540, 369]. The models vary in how they conceptualize each aspect of SRL and what skills they emphasize, but all focus on the same concept of learners leveraging cognitive and meta-cognitive techniques to analyze tasks, design strategies, and monitor performance. [497]. For example, Boekaerts's model emphasizes the importance of goals and suggests two pathways of SRL that students follow when attempting to reach their goals: a learning or mastery mode and a coping or well-being mode [75]. The first is followed when students perceive a task as congruent with their goals, while the second is pursued when the task is perceived as a threat to their well-being [75, 497]. In contrast, Winne and Hadwin's model places a strong emphasis on meta-cognitive skills [723, 497]. In addition to goal setting and performance monitoring, their model emphasizes processes in which students plan and adapt for the future by making changes in their motivations, beliefs, and strategies [723].

Hadwin et al. [273] also present a unique view on SRL through their model for co-regulated learning. Co-regulated learning is defined by the authors as a transitional processes in a learner's acquisition of SRL, during which members of a community share a common problem-solving plane, and SRL is gradually appropriated by the learner from a more experienced other. For example, a mother teaching a child to tie a shoelace might ask questions like "what do you know about how to connect those two laces?" and "How do you know when you have completed the first step properly?" This allows the child to focus on task enactment while an external force supports meta-cognitive engagement and regulatory control, thus easing the cognitive demands of completing the task. They emphasize that over time the student will acquire the regulatory skills of the other and will shift to a self-regulated learning arrangement. This concept of co-regulated learning is somewhat new, and there is therefore little work exploring it empirically, but I discuss it here as it is relevant in the context of games, which are sometimes learned in social situations, with more experienced others aiding newer players.

Literature demonstrated how SRL is directly connected to learning and academic performance. Zimmerman and Pons [757] found that high-performing students engaged the self-regulatory strategies of "seeking information", "keeping records and monitoring", "organizing and transforming", "seeking teacher assistance", "seeking peer assistance", "seeking adult assistance", "reviewing notes", and "reviewing text" far more than their low-performing counterparts. They also found that use of SRL was predictive of standardized test scores [757]. This later led to the development of scales for assessing SRL skill to better identify students who may need additional support [413]. In another example, Malmberg et al. [416] identified patterns in self, co, and socially-shared regulation. They discovered patterns in how successful students engaged the three processes over the course of their learning tasks, with co-regulated planning being the most frequent event and self-regulated planning being the least. They also found that groups engaged mostly in co-regulated planning and monitoring at the beginning of their collaboration, but as their collaboration proceeded from the starting phase to the intermediate phase, these decreased while task execution and socially-shared planning increased. At the end of the collaboration, co-regulated planning and monitoring increased again [416].

SRL has also demonstrated relevance outside of academic contexts. Previous work found that athletes leverage a variety of SRL processes when engaging with their sport, including task-understanding and goal setting, and that these can even transfer to academic contexts [434, 433]. Other work demonstrated the impact of SRL skills on athletic performance. Zimmerman and Kitsantas specifically investigated the impact of goal type on performance [754]. They found that those with process goals surpassed the dart throwing proficiency of those with outcome goals and that those whose goals shifted from one to the other further exceeded those with process. They conclude that beginning with process goals and then shifting to outcome goals at the point of automaticity is the key to successful skill gain whereas premature shifting to outcome goals is detrimental [754].

2.3.2 The Cyclical Phase Model

In the context of this dissertation work, I choose, specifically, to examine SRL in the context of the Cyclical Phase Model of SRL. The Cyclical Phase Model (CPM) of Self Regulated Learning is a foundational model of SRL proposed by Zimmerman [750, 756]. Building on Zimmerman's earlier models of SRL [757, 754, 413], The Cyclical Phase Model organizes SRL processes into three phases: forethought, performance, and self-reflection [750, 753]. An overview of this model can be seen in Figure 2.1. The forethought phase includes processes such as analyzing the task, setting goals, and planning how to reach them. The performance phase encompasses execution of the task and progress monitoring along with strategies to maintain engagement and motivation. The self-reflection phase encompasses the processes by which the learner assesses how they performed the task [750, 753, 497].

CPM has demonstrated relevance across various domains, including academic contexts [48, 456, 416, 385] and athletics [752, 755]. In academics, the skills associated with CPM have demonstrated direct correlations with performance. For example, Barnard et al. [48] sorted students in an online learning environment into profiles based on how they invoked SRL skills across the three phases of CPM. They found that students who invoked CPM strategies and skills had significantly higher GPAs than those students who either did not invoke any skills or who invoked them in disorganized and inconsistent manners [48]. In another example, Min and Foon [456] identified connections between the three phases of CPM and student engagement in a massive online open classroom (MOOC). Specifically, they found that emotional engagement mainly impacted the forethought phase (e.g. interest to learn the MOOC) and the self-reflection

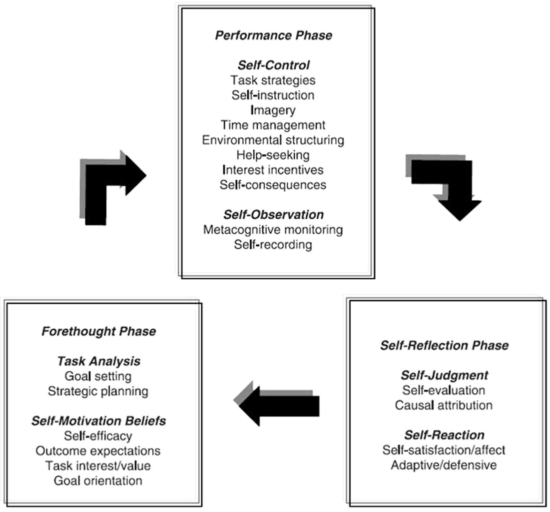


Figure 2.1: The Cyclical Phase Model of Self Regulated Learning from [756].

phase (e.g. enjoyment of learning). Cognitive engagement focused on setting the goals for learning and understanding, valuing the learning according to the learning outcome expectation, monitoring comprehension on the contents, and applying meta-cognitive strategies.

In athletic contexts, previous work discovered that execution of CPM processes are directly correlated with skill level and performance. Cleary and Zimmerman found that expert basketball players set more specific goals and technique-oriented strategies during the forethought phase and more often attributed failure to faulty technique during the self-reflective phase than non-expert or novice players [131]. Similarly, Kitsantas and Zimmerman [343] used the Cyclical Phase Model to study differences in SRL between expert, non-expert, and novice volleyball players. The results, again, found that experts set better goals and had better planning during the forethought phase, better strategy use and self-monitoring during the performance phase, and better evaluations, attributions and adaptations during the self-reflection phase than either non-experts or novices [343].

Other studies have also examined the impact of SRL on athletic performance. Cleary and Zimmerman used the Cyclical Phase Model in a study that examined the impact of the additive effects of self regulation training in forethought, performance, and self reflection processes on basketball free-throws [132]. They found that those who practiced all three phases of SRL had a significantly better shooting performance than those who only practiced one phase or none, indicating that SRL had a significant impact on overall performance [132].

2.3.3 Supporting Self-Regulated Learning and the Cyclical Phase Model through Prompts

While SRL skills can be taught through instruction of training, one of the most common approaches for promoting SRL is actually through prompts [44, 314]. Prompts are different from instruction as they do not teach new information. Instead, they support recall and execution of knowledge [45, 44].

In some scenarios, these prompts ask students to define goals for themselves [70, 588], thus supporting the forethought phase of CPM. For example, Graham et al. [259] found that prompting them to set goals helped students with learning disabilities write better compositions. More recently, Colthrope et al. [136] found that prompting pharmacy students to engage in higher-quality goal setting during the forethought phase could increase their learning outcomes. In relation to these findings, Handoko et al. [275] also found that goal setting had a direct influence on course completion in MOOCS, arguing for the increased inclusion of goal setting prompts in learning environments.

In other scenarios, these prompts are designed to trigger self-monitoring skills, thus supporting the forethought phase of CPM. While less commonly discussed in the literature in comparison to goal setting and reflection prompts, self-monitoring prompts have demonstrated the ability to improve learners' performance [332]. For example, Kauffman et al. [333] prompted students to take notes and collect information from a set of online learning materials, but one half of the sample was also prompted to self-monitor. The self-monitoring prompts were brief statements asking students to review their notes before moving on to the next activity. Results found that the self-monitoring prompts had a significant positive impact on note-taking quality and academic achievement [333].

There also exist prompts for SRL and CPM meant to trigger reflective processes, such as evaluation and adaptation, supporting the self-reflection phase of CPM. This is well illustrated by the work of Van der Boom et al. [682, 681], who found that elicited reflections improved students' performance in an online learning environment, especially when paired with educator feedback. In another example, Rakovic et al. [546] found that the quality of students' evaluations were predictive of adaptations in their learning process, which were in turn predictive of increased frequencies of desirable learning behaviors and higher exam scores.

Prompts, especially goal setting prompts, have also been used outside of academic contexts, in order to promote SRL and CPM processes in sports [714, 188]. For example, in an early study examining the interaction between goals and athletic performance, Wanlin et al. [710] found that prompting athletes, in their case, speed skaters, to set long and short term goals during practice resulted in improvements in performance. More recent work has continued to examine the role and impact of goals, such as that of McCarthy et al. [435], who found that goal setting may have an impact on positive affect, and therefore performance, in athletes. There is also the work of Blijlevens [66], who found that goals were critical to directing and orienting the efforts of gymnasts across skill and experience levels. Some work has also examined reflection support, such as that of Chow and Luzzeri [122], who developed a tool for reflectively examining athletic performances.

2.3.4 Supporting SRL and CPM Computationally through OLMs

While prompting is the classic approach to supporting and promoting SRL, previous work has also demonstrated that SRL can be elicited through interaction with data, typically made available to learners through Open Learner Models [294]. A learner model is "a machine's representation of the learner". More specifically, learner models, generally, use data regarding a learner's observable actions within an educational environment to generate a representation of that learner's knowledge. An Open Learner

Model is one that is revealed or presented, through some means, to the learner [73, 5]. OLMs implement a variety of features that elicit SLR processes across the three phases of CPM, which are reminiscent of the prompts discussed in the previous section. In the forethought phase, OLMs aim to help students set realistic goals and plan how they will achieve those goals. For example, Progressor [299] is an OLM that uses charts to visualize all of the topics in a course and uses color-coding to indicate which topics have been covered and how well the learner has grasped those topics. It is also a social OLM in that it allows students to see the anonymized charts of their classmates, allowing each student to see their own standing compared to others. The results of an evaluation found that the social approach helped students plan their learning strategies based on the information they had of the highest performers' performance in the course [299]. In the performance phase, OLMs help students monitor their progress on a task while executing learning strategies [294]. For example, INSPIREus [611] tracks student activity within a digital learning environment and presents summary evaluations of their performance, understanding of topics, and learning gains based on their actions. It also presents visual representations of learners' strategies, efforts, and working and learning styles, inferred from their interaction data. e-KERMIT [276] is an older OLM, that generates a model of learner knowledge and depicts student progress towards understanding course topics using progress bars. For both of these OLMs, user evaluations found that students often had a different impression of how much progress they were making than what the OLM showed them. Overall, students found that INSPIREus's interpretations of their learning strategy improved their self-awareness of their own behavior, and helped them monitor their actions better. Similarly, students who interacted with e-KERMIT found it helpful in allowing them to monitor and generally be more aware of their progress towards learning a topic.

In the self-reflection phase, OLMs aim to support reflection and error identification. Many OLMs leverage negotiation with the model to facilitate this. For example, NDLTutor [647], STyLE-OLM [162], and CALMSystem [334] all feature some form of negotiation. In the cases of both NDLTutor and STyLE-OLM, the model generates a representation of what the student knows based on their performance on assignments and assessments. Both OLMs also feature a means by which the student's own understanding of their knowledge can be input. In the case of NDLTutor, the student is prompted to provide a self-assessment at set points which will prompt a negotiation if there is a discrepancy between the self-assessment and the model. In the case of STyLE-OLM the student can browse the model through a GUI and initiate a discussion if they see something they disagree with [162, 647]. In both cases, the negotiation is a means by which the student can either convince the system it is wrong, and prove themselves, or come to understand that the system is correct. CALMSystem [334] has a similar set-up, but leverages a chat-bot to make interaction with the model feel more natural, whereas the previous two use text-based natural language communication. Evaluations of these systems found that being able to negotiate with the model supported reflection and resulted in fewer discrepancies between the model and the student's own understanding of their knowledge [647, 162, 334].

2.4 Games and Learning

2.4.1 Learning Complex Gameplay

Games literature has long since recognized learning as critical to successful gameplay and even suggested that games themselves can be vehicles for learning [239]. Perhaps the most influential work on learning gameplay is the work of James Paul Gee [238, 239, 604], who argued that games were environments for learning based on, among other things, how they require and encourage problem solving and provide feedback at appropriate times. The study of learning in games, however, is rather complicated. Much of the work on learning in games focuses on educational games or leveraging commercial games in academic contexts and is interested in the extent to which the game can teach the player the academic content [480]. Learning gameplay itself, in contrast, is notably under-explored in existing literature.

Despite expertise being recognized as a key component of gameplay [178], studies of expertise in complex games are often more interested in the skills possessed by expert players, rather than how those skills are gained. Nowhere is this more apparent than esports research, which often discusses and formalizes the skills experts use to succeed in gameplay [377, 207, 164, 544] but rarely examines how players gain these skills [508, 521, 287]. Outside of esports, learning gameplay is most often studied in relation to overcoming failure [71, 312, 313, 311] or in the context of community learning [487, 472]. In the former, literature discusses how failing to progress in a video game can lead to learning, as it forces players to problem solve and re-engage with the mechanics of the game in order to overcome the impasse [71]. In the latter, literature discusses how interaction with a community can promote learning through explicit guidance and social motivation [487].

While informative and thought provoking, this work fails to articulate specific, generalizable processes of learning gameplay that can be facilitated or supported through external interventions. Further, the work discussed above does not leverage formal learning theories, and specifically does not consider learning in terms of SRL or CPM.

2.4.2 Self-Regulated Learning in Games

In the context of digital games, SRL is notably under-studied, and much of the existing work focuses almost entirely on educational games [480, 481]. For example, [571] generated SRL scores for students who played the educational game *Crystal Island* [567] based on their responses to a reflective prompt. They found that SRL scores were significantly predictive of post-test learning gains and that high-SRL students appeared to make more use of the in-game curricular resources than low-SRL students and reported more immersion, interest, and enjoyment [571].

Similar to the work in the learning sciences, the study of SRL in digital games often evaluates the impact of prompts [480]. For example, [493] added a self-explanation prompt, which encouraged self-reflection processes, to an educational math game and found that students who responded to the prompts tended to have higher mean post-test scores than those who did not. Similarly, [214] found that adding prompts to a game that taught electrical circuits significantly increased student performance. Several studies have also examined the impact of different kinds of goals on performance in game-based learning. For example, [369] examined the impact of type and specificity of goals in a game-based learning environment that taught buoyancy concepts. Their results found that non-specific problem-solving goals yielded substantially more frequent strategy use from learners, but that this was not the case when the goals were learning goals [369]. [211] examined a similar question, but in the context of educational game design through Scratch [556]. Their results demonstrated that students with non-specific goals outperformed those with more specific goals and that students with structuring scaffolds demonstrated worse SRL [211].

While all of this work demonstrates the role that SRL can and does play in games, it focuses entirely on educational games and, in most cases, on the impact SRL has on players' learning of the educational content [214, 493, 211, 369]. In contrast, there is currently little work that examines the role that SRL plays in learning the skills and mechanics involved in playing a game. [88] provide one of the only examples of work that examines SRL skills in relation to gameplay performance itself. In their study, they investigated the impact of unrealistic performance goals on player performance in a first-person-shooter game [88]. They found that those whose performance fell short of their goal would perform significantly worse in subsequent levels than those whose performance more closely matched their goal. Further, they found that this was significantly more common for those players with high video-game self-efficacy [88]. This suggests that eliciting and supporting SRL skills in the context of learning complex gameplay can greatly benefit players who are trying to learn and improve at gameplay.

2.5 Summary

In the context of this dissertation work, I explore the potential to support SRL in the context of learning complex gameplay. While there exist many theories of learning [590], SRL, and specifically CPM, lends itself especially well to the context of complex games, as it formalizes the processes by which a student can learn on their own [750]. Despite this, SRL and CPM are currently understudied in games, with most existing work focusing on their relevance in the context of educational games [480]. This, compounded with the fact that there is little research on how players learn complex gameplay, means I know little about the role of SRL in learning to play and excel at complex games.

Further, I explore the potential to support self-regulated learning of complex gameplay through computational tools. I choose this approach for three reasons. First, related work demonstrated how computational tools, such as OLMs, can prompt SRL processes and elicit SRL skills to support learning [294]. Second, computational tools are already common in domain of complex games, especially esports, where numerous assistants are commercially available [601, 459, 68]. Third, computational tools can be interacted with in the absence of an educator or support network, making them well suited to support players who do not have access to such resources. By better supporting SRL and CPM in such a manner, we can make complex gameplay and its proven benefits more accessible to players and more effective in serious domains. In the following chapters I will discuss each of the research thrusts in this dissertation work in detail.

Part II

Studies of Self-Regulated Learning in Complex Games

In this first thrust of my work I discuss three formative studies I conducted that explored learning, self-regulated learning, and co-regulated learning in esports, recognizing esports as one of the most complex examples of a complex game. This work sets the stage for my later research on computational assistance by providing a foundational understanding of how learning occurs upon which I later build with more detailed explorations of how to better support that learning.

Chapter 3

Learning in Complex Games

The work discussed in this chapter is based on a paper originally published in CHI 2022 [1] [349].

3.1 Understanding how Players Learn Games

The first research thrust in my dissertation work seeks to answer RQ1: "How do players engage self-regulated learning skills in the context of learning and improving at play?" I explored this topic within the domain of esports, complex games that encourage players to learn and improve so they can succeed in ranked leagues and formal competitions. Further, esports are high-stakes, fast-paced, dynamic environments where players are forced to make choices with incomplete information and findings derived from such an environment may generalize to other, similar, environments, such as disaster response.

Expertise, gained through learning, is critically important to success in games, especially in esport games, as demonstrated by Eaton et al. [178, 177], who found that a player's experience in a particular role or with a particular character can have a direct

[1]Kleinman, E., Shergadwala, M. N., & Seif El-Nasr, M. (2022, April). Kills, deaths, and (computational) assists: Identifying opportunities for computational support in esport learning. In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (pp. 1-13). This was collaborative research led by me but could not have been possible without the continued support and input of Murtuza and Magy.

and significant influence on an esport team's chances of winning. As such, previous work sought to formalize the skills [544] and knowledge [207, 164, 377] that expert players possess and how they manage them in-game. For example, Fanfarelli [207] discussed how expert-level *Overwatch* players are able to anticipate what is to come and communicate openly and effectively with teammates in order to ensure survival and maintain good in-game awareness.

The work discussed in the previous paragraph presents comprehensive overviews of what skills are required to play at the expert level, but does not explore the means by which those skills are gained. Literature that looks more closely at learning and skill gain is relatively uncommon in comparison. One such example comes from Hesketh et al. [287], who conducted a grounded theory interview study examining how team-vs-team esports are learned. Their preliminary results found that players engage several learning processes including identifying skill gaps and applying existing skills in new scenarios. This work, however, is still in its preliminary stages. Another example comes from Pluss et al. [521], who explored the impact of practice quality on tournament performance. However, this work focuses on practice exclusively and does not seek to create a more general understanding of all the strategies that players engage in when pursuing skill gain. Further, neither of the above studies examined learning from the perspective of SRL. As discussed in the previous chapter, the study of SRL in games has focused almost exclusively on educational games [480]. As a result of being more uncommon as a research topic, the details of how players gain skill and move up in expertise are largely unknown.

In contrast, skill gain, and the methods employed in its pursuit, is a more popular topic in the literature surrounding traditional sports. One prominent topic in this literature is the impact of learning processes, such as goal setting, on performance and improvement in athletics [714, 188]. For example, in an early study examining the interaction between goals and athletic performance, Wanlin et al. [710] found that prompting athletes, in their case, speed skaters, to set long and short term goals during practice resulted in improvements in performance. More recent work has continued to examine the role and impact of goals, such as that of McCarthy et al. [435], who found that goal setting may have an impact on positive affect, and therefore performance, in athletes. There is also the work of Blijlevens [66], who found that goals were critical to directing and orienting the efforts of gymnasts across skill and experience levels.

Goals are not the only component of skill gain and improvement discussed in the sports literature, however. Social interaction, with both teammates and coaches, is also recognized as critically important. This is well demonstrated by Fahmi et al. [200], who found that a coach's leadership can affect teamwork and motivation, both of which directly impact performance. And of course, a notable amount of work in the sports literature also looked at the impact of practice itself [409, 283]. In one such example, Baker et al. [38] examined the correlation between the amount of sport-specific practice and expert decision-making. Their findings suggested a direct correlation between the amount of sport-specific practice and expertise, but that less sport-specific practice is needed to achieve expert-level decision-making if the athlete possesses prior experience in other sports [38]. This knowledge is valuable to the domain of sports as it allows coaches and administrators to make informed decisions about recruitment and training and can also help athletes overcome obstacles through targeted interventions.

This interest in skill gain and learning is, however, still new and largely understudied in the domain of esports, and thus, it is relatively unclear the extent to which existing findings from sports literature transfer across domains. While esports have been classified as sports [195], there are inherent differences from traditional, non-video-game-based athletics. Most notably, esports are played on a computer, and the games include user interfaces that provide players with game-state information that traditional athletes may not have access to, or at least not in the same way. Later in this dissertation, I will discuss results that suggested that the presence of these user interfaces may result in significant differences between how esports players and traditional athletes approach learning and improvement [345]. Further, the definition of esports currently accepted by the community is, essentially, "gaming that occurs in an organized, professional league" [221] suggesting that any game has the capacity to be an esport even if it possesses no mechanical resemblance to an actual sport. Thus, in the context of designing targeted computational support for learning esports, it is important to conduct a focused examination of how skill gain and learning occur in esports.

In this study in this first research thrust, I sought to address this gap and expand our understanding of how learning occurs in esports, specifically focusing on the role of SRL, specifically CPM, in learning and improvement.

3.2 Empirical Study of Learning Activities and Challenges in Complex Games

In this first study, I sought to address the gap discussed above, by examining how players learn and master esports games in order to identify opportunities to computationally support this process. Specifically, I sought to answer two research questions:

- What are esports players doing in order to learn and master gameplay?
- What challenges are esports players facing as they attempt to learn and master gameplay?

The first question specifically asks what players are doing, as in what activities they are leveraging, in order to learn to play their respective games at the highest levels. The second question build on the first by asking what challenges they face, which can help reveal further insights into what they believe to be effective methods for learning and what, specifically, they feel needs to be addressed. These questions were answered through an interview study.

3.2.1 Methods

3.2.1.1 Recruitment

I recruited 17 esports players from the UCSC student body and through convenience sampling. Based on preliminary reviews of the data, saturation was seen at 15 participants and recruitment was stopped at 17 after no new information emerged. Participants were required to be at least 18 years of age, located in the United States, and familiar with at least one esport game. For the purposes of this study, any games with multiplayer, competitive leagues, including fighting games, such as the *Smash Brothers* and *Street Fighter* franchises, and digital card games, such as *Hearthstone* and *Magic the Gathering Arena*, were considered esports. These features are a part of the definition of esports as presented by Formosa et al. [221].

3.2.1.2 Interview Protocol

One-on-one, semi-structured interviews were conducted between February and April 2021. The audio was recorded and transcribed manually afterward. All participants were asked all questions. If a participant's response was overly succinct, I would ask a follow-up question to try to get the participant to elaborate more. This was done twice. If the participant did not add anything after two attempts, I moved on to the next question. All interview transcriptions were reviewed by two researchers to ensure that even succinct responses answered the questions. Interviews lasted between 15 and 45 minutes, with longer interviews being due to more verbose participants. The average interview time was 28 minutes.

After receiving informed consent, I collected demographic data from the participant, including games played, years of experience, and hours played per week. Participants were allowed to discuss more than a single game during the interview, and many did, however years and hours were only collected for each participant's most played game. Age, gender, and ethnicity were not collected to avoid risks of identification and unintentional biasing of analysis and results.

I then asked a set of open-ended questions about the participant's goals as a player, what they do to try and achieve those goals and improve at play, and the challenges they face in that process. Following these questions, I asked the participant to think about a fictional computational support tool that could help them improve at their game (or games) of choice. This second part of the interview was effectively a

	Demographics
1.	What esports games do you regularly play or did you play when you were more active?
2.	How many hours a week do/did you play your most played game?
3.	How many years of experience do/did you have playing your most played game?

Table 3.1: The demographic questions used in the interview study. Interviews were semi-structured and follow up questions were asked as needed.

hypothetical design exercise in which I asked the participant several questions regarding the functionality of this fictional tool, how they would interact with it, and how it would improve their experience. Additional insights into players' practices and the challenges they faced could be inferred from the types of functionality they required from their fictional tools and their discussions of why they wished for such functions. Participants were told that the tool could provide any kind of computational support, be it AI, data visualization, or something else entirely. They were also told to think about the tool in a "sky is the limit" manner, not restricted by legal or technical limitations. Participants were allowed to skip any questions that they did not want to, or felt they could not answer. The full list of questions asked can be seen in Tables 3.1 through 3.3.

3.2.1.3 Data Analysis

Audio was transcribed into text using Microsoft Word, and the data was then segmented into lines based on how the text lined up in the editor. Lines with five or fewer words were combined with the previous line. I then worked with my co-authors to conduct iterative thematic analysis and line-by-line coding [237, 574] on the data. We, separately, performed open coding on 30% of the data set [99]. This was done in an iterative manner, once focusing on RQ 1 and a second time focusing on RQ 2. Each line was treated as a unit of analysis for coding. Responses that were in the vein of "yes", "no", or "I don't know" and contained no other insights, i.e. "oh yeah that would certainly be useful, I didn't even know if those kinds of things existed" (Participant 1)

	Goals, Practices, and Challenges
4.	When you play what are your desired outcomes?
5.	What objective do you seek to achieve as you continue to play the game over time?
6.	Can you describe your activities that you engage in to pursue your objectives?
7.	What about those activities do you find enjoyable and what do you think could be improved?
8.	What challenges do you face when engaging these activities as far as achieving your goals?
9.	Have you ever used any computational tools, such as spreadsheets, AI assistants, or data visualizations, to engage in these activities?
10.	When practicing, training, or just generally trying to get better at the game have you ever sought the aid of others (friends, coaches, people online, etc. . .)?
11.	When during gameplay did you seek help (in the moment, afterwards, etc...), and what prompted you to do so?
12.	For any of the activities that you discussed above, can you imagine any kind of computational tool, such as spreadsheets, AI assistants, or data visualizations, that could help you engage in the activity?

Table 3.2: The goals and practices questions used in the interview study. Interviews were semi-structured and follow up questions were asked as needed.

	Fictional Tool Design
13.	Now consider a situation where AI/ML can improve your experience engaging in the activities mentioned earlier or allow you to pursue your objectives in other ways. What kind of information does it take in?
14.	How does the tool take in information?
15.	What information does it provide to you?
16.	How much information does it provide you?
17.	Is it prescriptive ("you should do this")?
18.	Is it descriptive ("this is what is happening")?
19.	Is it evaluative ("this is how you did")?
20.	When, during gameplay, do you interact with the tool?
21.	Is the tool interacting with anyone else or just you?
22.	Are you able to give the tool feedback or instructions?
23.	How would this tool improve your experience?
24.	If such a tool existed, would you have concerns about using it?
25.	Wrapping up, can you quickly summarize what the tool does?
26.	And can you quickly summarize how you interact with it?
27.	And can you quickly restate how it would improve your experience?

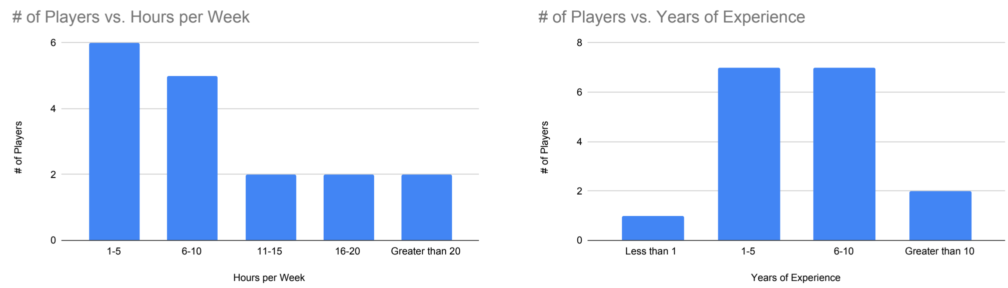
Table 3.3: The design exercise questions used in the interview study. Interviews were semi-structured and follow up questions were asked as needed.

were ignored. I worked with a collaborator to identify lines during the code derivation process that did not contain information relevant to one question or the other. These were discussed and, if we reached a consensus that the line was not relevant to the given question, not coded. This process resulted in approximately 2400 code-able lines for each research question.

After initial coding, we reconvened to compare and discuss our separate sets of themes, identify and collapse overlaps, and resolve disagreements. The final theme categories can be seen in Tables 3.5 and 3.6. The themes were then converted into codes and validated through inter-rater reliability using Cohen's Kappa [133]. The Kappa value for RQ 1 was .72 and the Kappa value for RQ 2 was .70, both indicating strong agreement [375]. I then applied the codes for both questions to the entire data set.

3.2.2 Results

The 17 participants recruited for the study had experience playing 14 different games. The breakdown of what games were mentioned can be seen in Table 3.4. Years of experience ranged from half a year to 21 years. Hours of play per week ranged from two to 60. These can be seen in Table 3.1.



(a) (b)

Figure 3.1: The years of experience and hours played per week for the participants for their most played game.

Game	Overwatch	Player Unknown's Battlegrounds	Smash Brothers Franchise	League of Legends	Counter-strike Global Offensive	Rocket League	Rainbow 6 Siege	Valorant	Team Fortress 2	Magic the Gathering Arena	Hearthstone	Marvel vs. Capcom Franchise	Street Fighter Franchise	Dota2
Participants	8	1	3	5	2	3	1	1	1	1	2	1	1	1

Table 3.4: The 14 games mentioned by the interview participants (top) and the number of participants who mentioned experience with each game (bottom). Note that participants could mention more than one game when asked what they had experience playing.

3.2.2.1 Activities

The thematic analysis for RQ 1 resulted in four activities. An overview can be seen in Table 3.5.

Activity	*Definition*
Practicing 24%	Improving one's skills and knowledge of the game through engagement with play.
Leveraging the Knowledge of Others 22%	Improving one's skills by referencing others' knowledge, experience, or presence.
Tracking Performance 29%	Improving one's skills by tracking one's performance over time.
Reflecting on Gameplay and Setting Goals For the Future 25%	Improving one's skills by reviewing past gameplay and setting goals to work towards

Table 3.5: The thematic analysis for RQ 1 revealed four types of activities players engage when trying to learn and master an esport game. Under the activity name is the percentage of total code applications for RQ 1 that were the given activity.

Activity 1: Practicing: Practicing refers to the act of playing the game and participant responses highlighted the variety of approaches that players engaged when approaching practice. While some simply mentioned repeatedly or constantly playing the game, i.e. "I guess mostly just doing matches" (Participant 2), others specifically mentioned using practice modes, dedicated tools, or AI opponents for more relaxed training or targeted drills. For example: "to train getting better at, like, sniper shots from far away, there's a map that you can download that has a whole bunch of training bots, so you can train hitting moving targets" (Participant 11) and "when I'm actually playing with the AIs, it kind of relieves that pressure and kind of just allows me to, like, practice without having to really think about winning or losing" (Participant 2).

Participants suggested that practice was an opportunity to build a stronger game sense, also known as a strong understanding of the mechanics and rules of play [207]. For example, participant 15 described their overall goal for practicing as "I hope, in Overwatch, to be able to have a better understanding of the Maps and to know where I'm positioned". In another example, participant 8 discussed developing game sense at a more granular, mechanical level: "you're basically playing soccer with the car and the manipulation of the car...those small things are actually incredibly important to winning matches and it's something I'm not very good at, so as I play I try to just, like, try stuff". The practicing code saw 579 individual applications.

Activity 2: Leveraging the Knowledge of Others: The second activity captures the role that other players play in an individual's learning process. Several participants mentioned playing with training partners, or a specifically selected other player who is recognized as having the ability and knowledge necessary to help test skills and guide improvement. In such scenarios, the knowledge of others typically took the form of personalized advice delivered in real-time. For example "being able to have conversations with [training partner] about the game, especially while we're playing, and figure out what I'm doing wrong or what I don't understand about a particular match-up or something has been extremely useful" (Participant 14).

Participants, especially those who played team-based games, also mentioned getting similar guidance from friends, who were often also teammates, who were at higher skill levels i.e. "[my teammates] give me some advice for how do I play...because I'm good at AD Carry [2] in this game, so I uh, they can advise me to do more practice or to eliminate some minions accurately or to improve my consciousness of the teamwork" (Participant 4). While participant 15 was the only one to explicitly mention working with a coach personally: "we have a coach on our team which works really well he's great", other participants did sometimes mention coaches as a resource available to players i.e. "there are people that you can get to like coach you and sometimes I have friends who do that and they'll give you advice, sort of, when you screw up, and that could help quickly, like, rectify mistakes" (Participant 8).

The knowledge of others could also be accessed in a less direct or personalized manner. For example, instead of getting personalized advice from a known other, some participants discussed watching others play, either in person or over streams i.e. "watching other content and learning almost by osmosis, oh, like, uh 'good streamer does good activity, therefore, I do good activity I do good'...looking up specifically informative videos and then otherwise just casually watching streams, watching professional play, especially analytical commentators about why certain things are working, why things

[2] A character class in League of Legends focused on powerful attacks.

are not, is most of my practice" (Participant 12). Others mentioned reading forums or guides, for example: "the main thing that I have to keep in mind is my placing and that's what, like, every time I read guides and, like, took that information in and actually placed myself well, that's when I did well in the game" (Participant 1). Participant 14 even mentioned a preference for guides over videos, citing that the latter typically goes off-topic: "I love watching people who are much better than you play and getting to, you know, when they're giving their thoughts and opinions I'm able to learn...but when they're very silent or they're, you know, talking about something not related to the game and just doing very strange plays it's hard for me to learn from that because I don't know why they're making the decisions they are making so kind of...articles are very nice because they're articles written and directed about the content." The code for this activity saw 520 individual applications making it the least frequent category.

Activity 3: Tracking Performance: The third activity captures all of the different ways that players are tracking their progress, including level, changes in performance, and overall score, over time. Some participants discussed examples of how the games themselves provided metrics or interfaces that they could use to track their performance. For example: "having the rank, the competitive rank, go up to see that my skill is reflected in the game" (Participant 6) and "at the end of each match, like, [it tells] you, like, how many hits you got or how many lives you had left" (Participant 13).

In other instances, participants mentioned using external tools that aggregate and archive gameplay stats over time. For example: "I used to follow a lot of my stats in Dotabuff.com. I used to keep an eye on, like, my win rate and which hero am I playing most of the time, like, frequently, and which hero I haven't played and I missed and I need to work on them" (Participant 17) and "There's a lot of tools online for league of legends that are essentially spreadsheets but there were, like, guides for the characters, like, on OP.GG or, like, Mobafire" (Participant 8). Some participants also discussed how they created their own records or tools to track their performance, for example: "I've kept track of, like, my win streak and stuff like that or if I'm playing in tournaments and I'm going up against people that I've gone up against before and I've beaten at one point and I have lost to them at other points I tried to keep track of, like, what character were they using and what character I was using." (Participant 13).

Some participants also discussed creative approaches to tracking performance, such as using in-game achievements as a metric when no other way of measuring performance was available. For example: "I just use, like, the achievement bars in the game itself to be, like, OK, like, say I have to win 10 Maps in one specific area in Counter-strike...to do that I'll play it 10 times and I'll try to keep getting better and just try to unlock the achievements" (Participant 5). The performance code saw 696 individual applications across, making it the most frequent activity.

Activity 4: Reflecting on Gameplay and Setting Goals for the Future: The final activity for RQ 1 is also the one most explicitly related to SRL. Specifically, this activity revolved around identifying mistakes that can be addressed in future play through review of and reflection on gameplay, as well as setting goals to work towards. As discussed earlier, under the theory of SRL, learners use both skills to plan and guide their learning [750]. With regards to reflection on gameplay, participants discussed that they typically engaged in post-play review with the intention of identifying weak spots to work on in the future. For example: "Uh I, sometimes, I watch my records, my competitions with others, like I always play my record and find my weakness and strength" (Participant 4) and "being told after like 'great, so, do you remember this part? This is where, like, you could have done something better' and just like thinking back on it" (Participant 15).

Regarding goal setting for future play, participants often discussed goals that involved reaching specific, trackable, in-game milestones. For example: "I just started playing ranked for the first time. I just played very casually for the past year or so that I've been playing and getting something like top 50% of players doesn't seem like an unreasonable goal" (Participant 12) and "well I recently hit one of the objectives with Hearthstone which was to actually hit legend and specifically hit legend with a deck that I built" (Participant 16). The reflecting code saw 595 individual applications.

3.2.2.2 Challenges

The thematic analysis for RQ 2 resulted in four general challenges. These can be seen in Table 3.6.

Challenge	***Definition***
Coordinating and Collaborating with Teammates 18%	Difficulties players face when trying to work together with others.
Knowing what to do Next 20%	Difficulties players face when trying to make the correct decision.
Tracking Game State 22%	Difficulties players face when attempting to keep track of the context of the game.
Tracking Skill and Improvement 40%	Difficulties players face when attempting to identify errors and evaluate performance.

Table 3.6: The thematic analysis for RQ 2 revealed four general challenges. Under the challenge name is the percentage of total code applications for RQ 2 that were the given challenge.

Challenge 1: Coordinating and Collaborating with Teammates: Participants revealed that, in the context of team-based esports, it is rarely easy to work with other players, and this trend is encompassed by this first theme. Those who played on teams discussed challenges related to communication and keeping track of teammates. For example: “I try to use VC [voice chat], like, all the time to try to communicate my messages across. The thing is that sometimes people don’t respond” (Participant 1) and “it could suggest, like, which teammate is, like, doing the most damage so I could, like, maybe, you know, focus on doing a little bit more [healing]” (Participant 9, describing a computational tool that could tell them which teammate was doing the most damage in-game). These challenges were sometimes exacerbated due to toxicity in the games’ communities, as demonstrated by participant 12: “it’s certainly unhealthy for the game when you can look up your current teammates right now and realize ’oh this guy is not playing his best champion he’s griefing me I’m going to flame him during the game’”.

Even in non-team-based esports, this challenge is a pressing concern, as interaction with others, in some capacity, be they teammates or opponents, is still relatively normal. Participants who played such games often discussed the challenges of finding someone to play with. This was especially concerning for those who discussed the importance of working with training partners, a theme discussed further in the previous section. For example: "my training partner growing up, and even into my adult years, has been my younger brother; and he's in a different part of the country than me and so we only really get to practice together, you know, before the pandemic, like, a couple of times a year" (Participant 14) and "I'm just playing with my younger siblings...it's, like, it's not very enjoyable just, like, beating them over and over again, you know, and it's, like, not fun for them either" (Participant 13). This code saw 435 individual applications, making it the least applied code for RQ 2.

Challenge 2: Knowing What to do Next: This challenge relates specifically to decision-making, a critical, and evidently challenging, element of the gameplay experience. Participants primarily discussed this in the context of real-time decision-making, articulating how difficult it could be to know what the best course of action in any given situation was and how, often, there was so much happening so fast that they did not have enough time to think. For example: "even when you're playing with a 5 stack and you're in [voice chat] with everybody and get to talk to everybody, you don't know what to do and, sure you have a player that hit diamond once in his life and is now vaguely your shot-caller, but do you fight here? Can you win? Do we see where all their players are on the map? Should we be taking this objective? Should we be prepping for this objective?" (Participant 12). In some cases, this difficulty arose due to uncertainty in how to respond to unexpected circumstances, i.e. "'cause I tend to, if someone appears behind me, I kind of freeze and I'm like 'Oh my gosh what do I do?'" (Participant 5). These quotes highlight how this challenge relates to one of the key elements of complex games: multiple correct solutions or strategies, with players rarely knowing the objectively correct course of action, or if there even is one.

Interestingly, participants also showed a general aversion to a tool that would explicitly tell players what to do. For example: "I think I would prefer if it just gave me the information and then let me figure out what to do next because...or like just the raw information is better for providing it to, sort of, myself and to the team and having...and then being able to use that information to then decide what we do next" (Participant 6). In some cases, this was out of concern that too much advice would interfere with learning to make decisions on their own, resulting in any kind of assistance ultimately becoming a crutch that the player would be completely reliant on. For example: "I think that if it gives me too much information, I'll feel babied, but if it gives me enough that I can learn on my own, I think that's best" (Participant 8). Players also suggested that such an aversion was to preserve the sensation that they were in charge of their own gameplay, for example: "I don't like doing that, 'cause it makes it feel like it's out of my choice and therefore, like, that doesn't make me feel good, when I built that sort of deck so anything that I use or add on is to the extent of something that I could myself do, it's just easier with said app, so note taking and such" (Participant 16, while discussing a deck building assistant for Hearthstone). This category's code saw 464 individual applications.

Challenge 3: Tracking Game State: The third challenge relates to the player's ability to track the larger context of play by keeping track of everything that is happening in the game, a challenge, given that there are often many things happening at once and many other players to keep track of. In some cases, participants discussed difficulties with spatial awareness, related to trouble tracking where other players, objects, or objectives were located. For example: "It might be difficult to figure out where opponents and teammates are 'cause they're not...they're not always on your screen and they're not always marked in any way" (Participant 8) and "I feel like maybe they could probably add the heal pack locations, 'cause they're kinda discrete and maybe that is the point, but, you know, sometimes it's really hard, finding it when you really need it" (Participant 2).

In other cases, participants discussed difficulties tracking resources, such as the number of cards in a deck or one's own or an enemy's health. For example: "I would use an add-on to help keep track of different aspects of the game, so if things were added to their hands, you know, things were returned to their hands, what were they" (Participant 16) and "in terms of CSing, you usually want to, like, last hit the minion to get the gold, so you want your shot to kill the minion or else you won't get the gold and personally, when I was a new player, it was hard to gauge when a shot from my champion would kill that minion" (Participant 7).

Similar to the previous theme, however, participants also expressed an aversion to any tool that would give them any information that they would not otherwise have access to, expressing concerns that such support would constitute cheating. For example: "so if they see, like, a certain enemy coming around the flank or something happening that I don't know is happening and then the computer tells me that it's happening, there's no way that I should have been able to figure that out based off of what I have on my screen and based off what I know of the game" (Participant 11) and "so aim-bots could be anything from like bullets hitting the target when it shouldn't or it could also be already seeing where the enemies are without having to actually see them, so if you're behind a wall they can't see you but then you see an outline of where they are, so any of those, like, is cheating because you have an advantage over other players that is just unfair" (Participant 10). This code saw 516 individual applications.

Challenge 4: Tracking Skill and Improvement: The final challenge encompasses difficulties that players face when trying to understand their own talents as a player and how they compare to those around them. There were two primary ways that players felt this challenge, the first was in identifying errors in their gameplay and the second was in evaluating their performance as a whole.

Regarding error identification, participants expressed that it was often difficult to know that they were doing something wrong, which could lead to repeated failure without knowing why, i.e. "I guess playing repeatedly didn't work, 'cause, like, I was repeating my mistakes, kind of, but I didn't know I was repeating it" (Participant 3) and "if I'm in a game and I make like a major mistake, sometimes it's hard for me to recognize immediately what I've been doing wrong and, like, where I've been screwing up and that can be pretty frustrating for me, 'cause it can feel like I'm losing and I don't know why"

(Participant 8). Some participants also discussed troubles identifying mistakes that occurred at a very granular level, i.e. hitting a button at the wrong moment in the middle of a combo. For example: "in order to keep my combos going into, you know, at what point did I lose the string? When did things go wrong? Getting more frame data from that would be wonderful" (Participant 14). Players often expressed frustration at experiencing undesirable outcomes but not knowing what was causing them.

Regarding evaluation, players suggested that it was often difficult to know if someone was playing well, better than they had previously, or at a desired level of competence. At the same time, however, participants were a bit divided on whether or not they would want a tool to do this for them. Some participants expressed an interest in an objective evaluation or grading system that could provide a better understanding of the quality of their performance. For example: "I love rubrics, I would love if you were to tell me, like, 'because of your KD [kill/death] ratio, because of how much you healed and the amount of time that the match was, you get this grade' it would be so fun" (Participant 15). Others, however, discussed concerns about receiving inaccurate evaluations or that an evaluative tool may not know as much about the game as they do. For example: "my only context for evaluating metrics is stuff like what Riot currently does...where it's pretty well known that you can get an S and not do well in the game" (Participant 12) and "if you are at rank five and, rank one being the best rank the topmost rank. I remember [an app] telling me that, even though I am rank five, my game play was, like, rank six level and I was like 'please no don't tell me how to play this game' and I just deleted that bot" (Participant 17). This code saw 956 individual applications, making it the most frequent challenge by a notable margin.

3.3 Learning, Challenge, and the Cyclical Phase Model

In the context of the original work [349], these eight themes were used to distill six explicit implications for supporting learning in esports through computational tools. These were: *evaluations provide context for understanding tracked data, explanations* *may be necessary to promote trust, players require practice partners that play like real people, real-time data tracking can act as an alternative to proper communication, state tracking can improve situational awareness and game sense, computational tools can use scaffolding and co-regulated learning to support decision making without becoming a crutch.* Within the flow of this thesis, however, I instead focus on what these findings mean as far as understanding learning, and understanding SRL, in the context of esports, one of the most complex genres of complex gaming that exists. If you wish to read more about the implications for tool development please refer back to the original paper [349].

Almost all of the themes relate to and can be understood as manifestations of some element of Zimmerman's model [756], with the fourth activity reported in the previous section (reflecting on gameplay and setting goals) being the most prominently related (reflection and goal setting are explicit processes that occur within the model, as discussed above). The activity of tracking performance over time and the challenge of tracking skill and improvement both relate to the self-evaluation phase of SRL, as it is through evaluation, attribution, and adaptation, the three key elements of this phase [132], that one is able to measure and track performance. Similarly, players face this challenge when they are not able to execute these processes, in other words, when it is explicitly difficult to evaluate their performance accurately and choose a correct attribution (cause of failure) or course of adaptation.

Similarly, the activity of practicing can be understood as related to the performance phase. It, essentially, provides players with the opportunity to, not only develop game sense but to become more comfortable with the skills of monitoring performance and adjusting strategies appropriately. In relation to this connection, however, the challenges of knowing what to do next and tracking game state can be seen as roadblocks to the execution of this phase of SRL. Specifically, knowing what to do next is connected to adjusting strategies by making informed decisions [756, 343], and facing this challenge suggests that players may be struggling to do so. Similarly, monitoring one's performance (a performance phase SRL skill) in-game is a part of monitoring the overall state of the game, and thus relates to the challenge of tracking game state, in that being unable to track the game state would interfere with the execution of this skill.

That players discussed activities and challenges related to elements of SRL and CPM in this way emphasizes their relevance, as theories, to the context of learning in complex gameplay and suggests that supporting these activities and addressing these challenges would, in turn, better facilitate SRL in the domain. Interestingly, however, one activity "leveraging the knowledge of others" and one challenge "coordinating and collaborating with teammates" suggest that learning within the domain of esports, and likely for all complex games, is not, strictly, a solitary experience and that there are, in fact, social components. These social components suggest the relevance of the related theory of Co-Regulated Learning (CoRL) [273, 272], which I discuss briefly above and expand on more in the next chapter.

3.4 Summary

In summary, this work examined and highlighted the general means by which esports players learn and the challenges they face in that process. By making connections between these practices, challenges, and the three phases of CPM, I suggest a connection between CPM and esports learning and propose that players of complex games, especially esports, are, in fact, leveraging SRL and CPM when trying to improve at play, even if they do not explicitly know what SRL is. With this proposal in mind, the next study in this thrust aims to build on this work by briefly digging into the social component of learning that was suggested by two of the resulting themes.

Chapter 4

Co-Regulated Learning Among Esports Teams

While the focus of this dissertation is primarily on solo learning via SRL, I have briefly laid the groundwork for future research on social learning, informed by the theory of Co-Regulated Learning [273], which I discuss further in this chapter, and prompted by the suggestion that social components play a part in the learning process, as implied by the results of the previous study. This work is currently under review at CHIPlay 2023 [1]

4.1 Co-Regulated Learning and Social Esports

Co-Regulated Learning (CoRL) is a theory of learning related to SRL. According to the theory of CoRL, SRL skills are appropriated from other more experienced individuals through a process of giving and receiving input throughout the execution of a task [273, 272]. Essentially, according to CoRL, a more experienced other can handle metacognitive aspects of a learning task, such as setting goals or monitoring progress, while the learner focuses on the actual completion of the task, as in the mechanical

[1]This research project was led by me but conducted in collaboration with researchers at Abeline Christian University: Dr. James Prather, Dr. Brent Reeves, Garrett B Powell. Further, this work would not have been possible without the assistance of Reza Habibi at UCSC and the input of my advisor, Magy.

steps needed to perform the task. This setup, CoRL argues, reduces the cognitive load of the learner as they focus on learning the step-by-step process by which the task is performed. Over time, as they gain skill and the steps needed to complete the task become more automated, the learner will begin to manage more of the metacognitive processes themselves, eventually switching to an SRL setup in which the other individual is no longer needed

To date, CoRL has not been explored in esports, and very little work has explored the phenomenon or expanded the theory in general. Metacognition, however, has seen a fair amount of attention in computing education, where self-regulation has long been seen as an important skill in the context of learning to code [530, 402]. Researchers originally called attention to this important aspect of computing education prior to 2010 [55, 470], and interest in it has exploded since then [689, 373, 421, 203, 691, 400, 401, 315, 534, 533]. The studies listed here draw from older theories of metacognition and self-regulation by Bandura [39], Flavell [218], Pintrich [515], and Zimmerman [751]. Application of newer theories, such as CoRL [273, 272] remain nascent.

Although a couple of abstracts proposing studies that utilize CoRL have appeared recently [620, 618], the first in-depth study only appeared in 2022 [531]. In this study, Prather et al. explored the reflections of students in an introductory programming course (n = 1,000) on their group study behaviors. They found that students exhibiting higher self-regulation knowledge and behaviors generally performed better, but higher co-regulation did not correlate with better performance. A qualitative exploration of the themes from these reflections revealed several that are relevant to the present study. The most common co-regulation theme was social help-seeking behavior. These students discussed how they asked their peers for help after attempting the problem on their own or asked to see solutions from friends after solving the problem to see how others did it. The second most-common theme related to group learning. Students indicated that they appreciated being able to bounce ideas off each other while solving programming problems. The third-most common theme was that of socially-shared regulation. This theme comprised elements of ensuring everyone in the group is on the same page, reminding each other of goals, and holding one another accountable. Finally, the fourth-most common theme was that of learning through teaching. While informative, and providing one of the first detailed, empirical explorations of how CoRL occurs in the wild, this work looks specifically at computing education rather than esports.

While there is work examining questions of teamwork and collaboration in esports, these rarely focus on learning. Instead, there is an interest in the elements of teamwork that impact or predict success [486, 4] such as how team members of different expertise levels work together [178, 177]. One example of such work is that of Musick et al. [473], who identified three major themes regarding how team cognition is perceived to be in relation to and an indicator of success in esports. These were: 1) shared awareness of dynamic game flow (essentially, being on the same page), 2) mutual understanding of skills and personality (essentially, everyone having the same understanding of everyone on the team's skill), and 3) sharing behavior in actual gameplay (essentially, how players choose who to share information with and what to share) [473].

Other work has similarly looked at the specific detriments to success through various means. For example, Zhang et al. [742] determined that players who frequently switched teams and teams that recently gained new players had lower overall performance in Counter-Strike: Global Offensive (CS:GO). Similarly, Parshakov et al. [503] found that CS:GO teams with low cultural diversity increased team performance but that experience and language diversity decreased it. Poulus et al. [527] found that players saw building and maintaining team dynamics as a challenge and that this was a detriment to success in esports, as team dynamics are a core component of successful team esport play [658]. As a result, Poulus et al.'s participants reported actively trying to prevent interpersonal conflict while also acknowledging that disagreements and communication breakdowns were inevitable [527].

Another element of teamwork that is studied in esports literature is communication, due to its importance to success [395, 226, 486, 4]. Tan et al. [656], for example, studied the connection between team communication, team cohesion, and success in ad hoc League of Legends (LoL) teams. By looking at correlations between communication sequences and team cohesion, Tan et al. [656] found that team cohesion had a connection to, and was in some ways predictive of, match outcome and team satisfaction. They also found differences in how low and high-cohesion teams communicated, with more cohesive teams being more likely to apologize, encourage each other, and agree to a suggestion [656]. Team cohesion is recognized as critical to success in traditional athletics [471, 107] and thus there is increased interest among practitioners in developing protocols for increasing cohesion among esports teams [649, 110].

While information about how teams work together to achieve success in one of the most complex gaming environments, the above work does little to help us understand how learning occurs as a social activity. Some existing work has discussed how information is passed among community or team members and how players may leverage the insights of experienced others to guide their own improvement [287, 225]. However, to my knowledge, there has yet to be a focused, empirical study of how players regulate each others' learning in the context of esports.

4.2 An Exploratory Study of Co-Regulated Learning Among Esports Teams

To develop an initial understanding of how players helped each other learn, informed by the theory of CoRL [273, 272], I conducted a semi-structured interview study that sought to answer the following question:

Based on these theories, this work sought to answer two research questions:

- What co-regulated learning themes emerge in the context of esports teams?

This work explored this question alongside questions regarding emotional co-regulation and revealed a number of prominent themes surrounding how input is given and received by teammates. For the purposes of this dissertation, I focus on the elements of this work related to CoRL and refer you to the original paper if you wish to read more about the study of emotions that occurred in parallel.

4.2.1 Methods

4.2.1.1 Interview Structure

Interviews were conducted in a one-on-one, in-person, semi-structured manner between October and December 2022 by a collaborator at ACU. Audio was recorded and transcribed afterward. After receiving informed consent, the researcher collected demographic information and then asked the questions seen in Table 4.1. Part one asked for general information about opinions and habits surrounding social play.

Parts two and three were informed by CoRL [273, 272]. As discussed previously, CoRL suggests that a learner focuses on mastering the physical skills of a task while the metacognitive aspects are handled by a more experienced other and eventually appropriates the metacognitive aspects as the physical skills are mastered, ultimately transitioning into a self-regulated arrangement. The other's handling of metacognitive skills manifests in the form of input given regarding evaluation, planning, monitoring, etc... To illustrate this, Hadwin and Oshige [272] provide the scenario of a mother teaching her child to tie shoes as an example. While the child manipulates the laces, the mother gives input in the form of statements like "what do you do next?" (planning) or "what did you do wrong?" (evaluation). The child can also take the initiative by asking, for example, what they should do. Focusing on these constructs, I developed the questions in sections two and three to target the phenomena of seeking and providing input. The exact questions were adapted from and built upon those used by Prather et al.'s study [531], discussed in the previous section, which also explored this facet of CoRL. The goal of these questions was to determine what metacognitive processes are co-regulated within the context of esports play.

The full study also contained interview questions targeting emotion that we do not expand on here as our focus is the question surrounding learning. All participants were asked all questions. James would ask follow-up questions or for elaboration if necessary. Participants were allowed to skip any question they did not wish to answer. University IRB approved the interview protocol. Interviews ranged from 14 minutes to

Construct	**Interview Questions**
Playing with Others	1. On a scale of 1 to 5 how often do you play, practice or train with others? 2. Would you like to provide further details regarding how often you train alone or with others? 3. What is your relationship with the other people you play with? 4. What do you typically discuss before, during, and after play? 5. What practice or training techniques do you find most effective when playing with others? 6. Do you prefer practicing or training with others or solo and why?
Seeking input	7. On a scale of 1 to 5 how often do you seek others' input on your gameplay? 8. What situations or circumstances prompt you to explicitly seek the input of others? 9. In relation to when you notice a mistake, especially in-game, when do you specifically seek out others' input? 10. Whose help do you seek and why? 11. Can you describe how they help you? 12. Can you describe any situations where you would go to community resources rather than known others? 13. In what situations or circumstances would you not seek the input of others if any exist?
Providing input	14. On a scale of 1 to 5 how often do you provide input to others? 15. Who do you provide input to and why? 16. Who initiates the interaction? 17. In what situations or circumstances do you provide input to others? 19. In what situations or circumstances would you not provide input to others?

Table 4.1: The Questions used in the interviews. Questions 1 and 7 used a Likert scale where 1 was the least and 5 was the most, participants were informed of this when asked the question. All other questions were open-answer.

41 minutes, with an average duration of 26 minutes.

4.2.1.2 Recruitment

14 participants were recruited via email and word of mouth from ACU. In order to participate, players had to be (1) at least 18 years old (2) Located in the United States (3) Able to communicate in spoken English (4) Play at least one esport. For the purposes of this study, any organized, competitive gaming was considered an esport, as defined by Formosa et al. [221].

4.2.1.3 Data Analysis

The data were analyzed via iterative thematic analysis by myself and a collaborator using a protocol similar to that discussed in the previous study [237, 574]. In an initial pass, we both, separately, reviewed a representative sub-set containing 30% of the data [99] and developed initial codes. The unit of analysis was a single utterance by a participant, which may have been an answer to a question presented in Table 4.1, a follow-up question, or a continuation of a previous thought interrupted by a comment from the researcher.

We then reconvened to discuss their independent code lists, collapse overlaps, and generate a combined code book. For reliability, we then coded a different 30% of the dataset [99] in order to measure agreement using Cohen's Kappa for Inter-rater Reliability (IRR) [133]. The resulting kappa value was .79, indicating strong agreement [375]. I then coded the entire dataset.

4.2.2 Results

4.2.2.1 Demographics

The demographic data for the participants can be seen in Figure 4.1. All participants identified as male and reported having friendly relationships with their teammates.

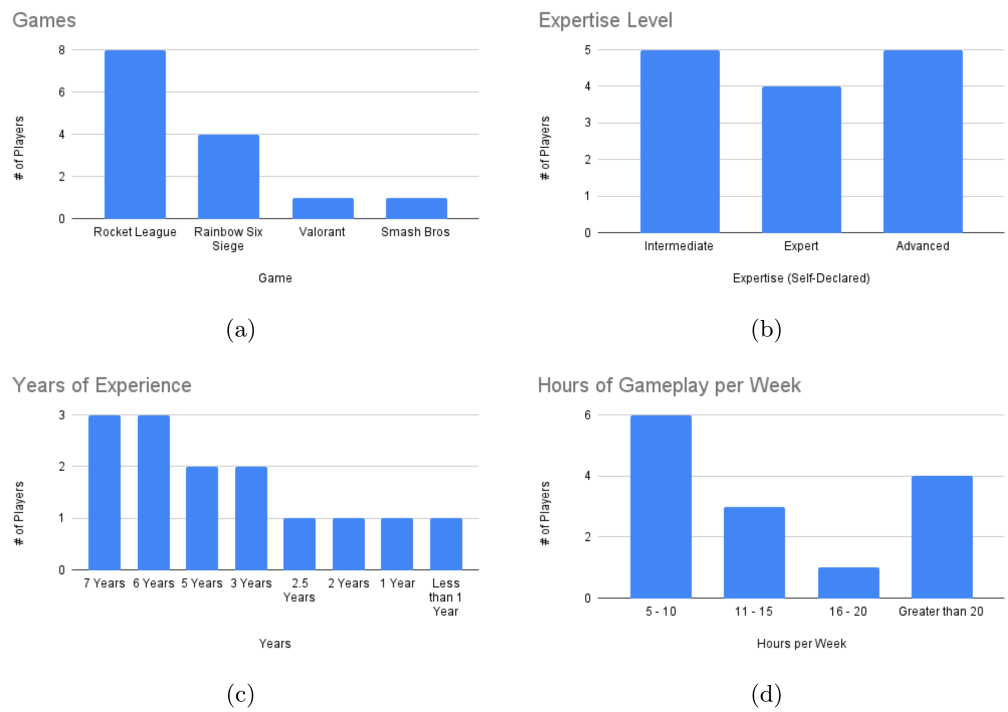
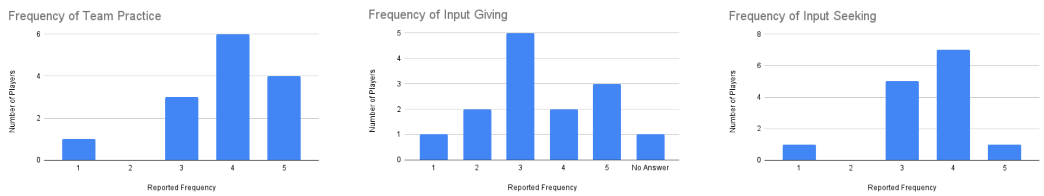


Figure 4.1: An overview of demographic data collected during the study.



(a) (b) (c)

Figure 4.2: An overview of event frequency data collected during the study.

4.2.2.2 Frequencies

The numerical data players reported for how often they play with others, give and receive input, and encounter conflicts are reported in Figure 4.2.

4.2.2.3 Co-Regulated Learning

An overview of the 18 CoRL themes and how often they appeared in the data set in terms of the number of total observations and the number of players who brought it up can be seen in Table 4.2.

Category 1: Relationship between Learning and Gameplay This category contains five themes related to how players described the relationship between learning and gameplay in the context of social and co-regulated learning.

Different learning goals warrant different approaches to play: All players discussed how they would choose to play with a group vs. solo based on what their overall learning goals were for the gameplay session. In general, playing with a group was better for developing team-related skills such as communication and obtaining an overall better sense of gameplay, while playing on one's own was better for drilling individual mechanics or learning new skills. For example: "I think practicing with other people is better. If you want to get better...playing in the team is better because you need to be able to communicate well and play well together and solo practice is just about improving your own mechanics" (Participant 1) and "a big part of Rocket League, especially, and Valorant, really, is getting in that solo time and getting the technique or the

# Obs.	*# Ps.*	*Theme*
		Relationship between Learning and Gameplay
59	14	Different learning goals warrant different approaches to play
30	14	Group reflection occurs after gameplay
24	12	In-game communication is focused on maintaining performance and achieving victory
18	14	Goals and strategies are established through group discussion pre-gameplay
4	3	The performance of the team is prioritized over the performance of the individual
		Goal Setting
10	8	Those with higher expertise or authority might help an individual set their personal goals
10	7	Teammates hold each other accountable for realizing goals
		Personal vs. Public Resources
19	11	Community resources can provide general information and instructions
7	5	Known others can provide personalized feedback and guidance
		Giving and Receiving Input
70	14	Input is expected to come from those with higher authority/skill/knowledge/experience
29	14	Other perspectives are recognized as critically important to learning
24	14	Recognition of failure, feelings of responsibility, or lack of knowledge will prompt input seeking
23	11	Input focuses on causes of failure and opportunities to handle the situation differently
22	13	Identification of a failure or repeated mistake will prompt input giving even if not sought
18	11	Input is not needed if the player is already aware of the situation
15	8	*Input should be tactful/constructive/clear*
9	7	Input requires an established relationship between the parties involved
5	4	*There is an expectation that input will be accepted*

Table 4.2: 18 themes were identified regarding how Co-Regulated Learning occurred among esports teams and these were organized into four categories. # of obs indicates how many times each theme was observed while # of pls indicates how many players discussed each theme.

different skills down by yourself, maybe improving your aim for Valorant or being able to do new tricks in Rocket League. But if you don't practice with your team...you're not gonna have the chemistry that you need for when you get into games and play together" (Participant 11). Notably, despite recognizing the benefits of both, players expressed a general preference towards group play and reported frequent group play as seen in Figure 4.2(a).

Goals and strategies are established through group discussion pre-gameplay: All players reported that, typically, prior to gameplay, the team would sit down together to discuss goals and strategies for that gameplay session. For example: "Besides winning and just maybe...'Hey, we need to try and rotate better'. We'll just say...not necessarily like...a number goal or something like that, but just 'Hey, this is our goal. We need to try and rotate better in these games'" (Participant 5, discussing goal setting prior to gameplay) and "sometimes...I'll bring up stuff like 'Hey, I just developed a new strategy for one of the sites we're playing make sure you take a look at that" (Participant 9, discussing strategizing before play).

In-game communication is focused on maintaining performance and achieving victory: 12 players emphasized that in-game conversation should be focused on ensuring that the game continued to run smoothly and that performance, morale, and chance of victory could be maintained. This meant that any conversation related to learning (such as identifying mistakes or giving input) did not occur during gameplay. For example: "I don't want to talk while we're still playing because it'll mess up the flow of communication. So generally...we try not to...say anything about mistakes until we're not actively communicating" (Participant 3) and "No, no, that's actually something that is not ideal, especially in teams...giving advice or correcting something right away in the middle of the game...phrasing it like 'Hey, you know, you could have done this'...it makes it much worse" (Participant 7).

Group reflection occurs after gameplay: In line with the previous two themes, all players suggested that learning occurred primarily after gameplay, a moment typically characterized by group reflective practices including evaluation of performance and suggestions for improvement, with an emphasis on getting others' input in the process. For example: "If we lose, we'll talk about what happened. We'll look at the game and see what we did wrong, what we did right...and if we win we're kind of happy about it" (Participant 10) and "There's a function in the game where we can review matches we've already played. So our team captain will go in on that. And then kind of just watch what we've been doing basically, and what we did do in that previous game. And then...he'll use that to tell us, I guess it's kind of like strategy, but like after the game, kind of for next time, I guess. But basically just still critiquing what we could have done better if we did lose or whatever" (Participant 13).

The performance of the team is prioritized over the performance of the individual: Three players implied that the performance of the team should be prioritized over the performance of the individual, going so far as to suggest that failure to do so was a sign of lower skill. For example: "When I was lower ranked, nobody wanted to talk. It was just like every man for himself and hope that you've got either a really good smurf on your team or somebody just has a really good game" (Participant 4, discussing how players would forgo proper communication at lower skill levels, preferring to focus on their own gameplay) and "So there was a very specific situation I was in and I was the last person alive. And...I think there were two enemies left on the opposing team. And basically what happened was...I was holding an angle and waiting because there was only one entry point to the room that they could have gotten in...but then for some reason, I decided to walk out and try to shoot them and I died because of that, and lost the match" (Participant 13, discussing how the team lost due to a move he attempted to make and how this was disliked by his teammates).

Category 2: Goal Setting This category contains two themes related to players' discussions of how goals are set and reached in the context of social and co-regulated learning.

Teammates hold each other accountable for realizing goals: Seven players discussed how other people would often step in to help them make progress towards and realize goals. Specifically, these participants reported that, often, they hold their teammates accountable, and their teammates hold them accountable, for realizing goals. These may be group goals, for which everyone is expected to do their part, or they may be individual goals that were shared with the group such that the group can check in on their progress. For example: "what we do is think of one thing that we each want to improve on for the session and then we, all three of us as a group, think of something that we want to improve on as a team" (Participant 1) and "we just try and remind ourselves on what we work on throughout the practice. And we help try and remind each other of that during the practice as well" (Participant 5).

Those with higher expertise or authority might help an individual set their personal goals: Eight players similarly suggested that other players may also help them set goals, typically someone of higher expertise or authority, such as a coach or more advanced player. For example: "Recently, [advanced player] has been coaching the B team with...specific training to follow, I know with [less advanced player] he's given [him] a specific set of training packs to go through" (Participant 3) and "One of our players, he is very new to the competitive scene so he's still getting used to the pacing and the intensity of playing at a competitive level. So I just talk to him and like 'Hey you need to work on your situational awareness and being able to process a lot of information while under pressure'" (Participant 9).

Category 3: Personal vs. Public Resources This category contains two themes that emerged from players' discussions of when they would ask a teammate for assistance vs. going to a community resource such as a forum or stream.

Community resources can provide general information and instructions: 11 players discussed how they use community resources to gain general information about play, or instructions about how to learn a new skill, instead of bothering teammates for this information. For example: "There's a lot of YouTubers out there that break down and analyze the game so watching them has helped me a lot so I recommend that" (Participant 9) and "Sometimes I'll just be like, 'okay, I need to fix this aspect of my game'. So, I'll go to YouTube and look up a specific mechanic that I messed up and be like, 'okay, how can I improve on this mechanic? How can I do it better?' I especially did that before I was in [team name] because I didn't really have anyone to go to so I had to go to YouTube. And sometimes when you ask someone for help...they'll literally give you a YouTube video" (Participant 11).

Known others can provide personalized feedback and guidance: At the same time, however, five players explicitly discussed how they would go to their teammates or coaches for input because it was more personalized and better suited to their specific problems. For example: "And if I don't know what to improve on then I seek help inside [team name]. Usually, they can point me in the right direction" (Participant 1) and "If I was trying to get specific help on...either a mechanic or a strategy in the game, I would probably just stick with my teammates and people associated with the team" (Participant 5).

Category 4: Giving and Receiving Input The largest category, this last grouping consists of nine themes that emerged from players' discussions of how they navigate the intricate process of giving and receiving input.

Other perspectives are recognized as critically important to learning: All players emphasized that an external perspective on one's gameplay could reveal new insights and that it was, therefore, important to get others' input on your gameplay. For example: "we'll watch what the captain sees mainly and we have to ask him like 'hey go to this round, check where I was, check who killed me, because I didn't see it'" (Participant 2, discussing the value of having his team captain watch his gameplay to help evaluate his decisions) and "Play styles are being introduced, like, by the community, almost every year, and...it would be impossible to get to a higher level without seeking out help on almost every aspect of the game" (Participant 12).

Recognition of failure, feelings of responsibility, or lack of knowledge will prompt input seeking: All players suggested that they were usually prompted to seek others' input by either their own recognition of failure or a lack of knowledge about how to handle a situation, possibly combined with feelings of responsibility. For example: "Mostly in a loss of a match or series. And especially...if it wasn't a close game. I'll definitely say 'hey, listen...something was up. I don't know what we were doing wrong.' Every once in a while if it was a close game, and I feel like I cost the loss...specifically, I'll get their input as well" (Participant 5) and "When I know I've messed up - like Rocket League I rotate slow - I know that I did something wrong in order to get back or something" (Participant 11, discussing when he asks his teammates for input).

Identification of a failure or repeated mistake will prompt input giving even if not sought: 13 players suggested that input does not have to be explicitly sought to be given and that a player may receive unsolicited advice if their teammates notice a prominent weakness or repeated mistakes. For example: "If I noticed them making the same mistake multiple times in a row, and maybe they don't realize it, like typically after you make a mistake, that gets you scored on, it will be like, 'Oh, why did I do that?' Or something like that. But if I see them make the same, like small mistake a few times and they don't mention it [I'll bring it up] just to bring it to light and that kind of thing" (Participant 6) and "Sometimes I need people to point it out because...I push too aggressively so I need somebody to tell me 'you need to stay back a little bit more'" (Participant 10).

Input is expected to come from those with higher authority/skill/knowledge/experience: While discussing how input was given and received, an explicit hierarchical structure came to light, with all participants discussing how the input should come exclusively from those recognized as having more skill, knowledge, or experience, or being in a place of authority, such as a coach or team captain. Some players even went on to suggest that not adhering to this hierarchy would result in conflicts among teammates. For example: "Usually [name]...because he's the captain of our team, so he knows...the best, like game knowledge of our team. And so he's the one I usually go to. Or if there's someone from like the [higher level] team, there's one of them there then I'll ask them because they know more than either me or [captain's name]" (Participant 3) and "Like if someone...like in Rocket League...someone's like, SSL (supersonic legend) and I'm like, champ, I'm like, 'yeah, you're gonna know better what...what I'm doing right and what I'm doing wrong.' So, yes, when they're better than me. I think sometimes when, which might be a fault of mine, but when people are on my level, I tend to be like, 'No, I don't want to get your opinion on this. Like, I know what I'm doing just as much as you do'" (Participant 11).

Input focuses on causes of failure and opportunities to handle the situation differently: 11 players specifically discussed how input received and given often focuses on helping identify and understand mistakes and how to avoid them in the future. For example: "There were things where I would do something, and I'm like, 'so this only works because I'm playing against people who are lower skill level than me'. I check to see if like, is it something that people are going to do in higher tiers, so I would ask [coach] some questions that he'd...like 'No, yeah, what you did was good, and you're spot on it only works because you're playing against a bunch of like, gold and silver players. Here's what you actually should have done or like here's...how it usually goes in higher tiers' (Participant 4) and "There are times where they'll just make a mistake. Just a simple mistake...missed the ball or something. And they'll ask, like, 'What should I have done there?'" (Participant 12).

Input is not needed if the player is already aware of the situation: 11 players also suggested that there was no need to seek or give input if the player already knew what they were doing wrong or what they needed to work on. Some players even suggested that giving input in a situation like this could cause tension or conflict. For example: "If it's something that they, like, clearly, already know what they did wrong. I don't think it is really helpful to say that again, because then they might just snap at me or something" (Participant 3) and "If it's a recurring mistake that I make, quite often I would say, you know, 'I'm working on this. I don't need any more input. I'm already very aware of the mistake'" (Participant 5).

Input requires an established relationship between the parties involved: Seven players suggested that seeking and giving input required an established relationship between the parties involved, i.e., two teammates, a player and a coach, etc. For example: "I think if I had a question that, like, no one on the [school] Rocket League team could answer...I think the thing I would go to is...YouTube videos...I don't think I would like ask the question to someone outside of [school] Rocket League" (Participant 3) and

"Specifically...my two teammates. I wouldn't necessarily give any input to the other team, but my two teammates I would" (Participant 5).

There is an expectation that input will be accepted: Four players implied that there existed an expectation that input given would be accepted, going on to suggest that refusal to accept the input could lead to tension or conflict. For example: "If I've already talked about it...and then they're still making the same mistakes like the next practice? It's like, okay, why did I even bother saying anything if you're not going to apply what I said?" (Participant 4) and "Especially for one player...he does not accept input very well because he thinks he doesn't need any" (Participant 7).

Input should be tactful/constructive/clear: In addition to discussing the details of when input was given and what it was about, eight players specifically emphasized the importance of tone and clarity when one gives input. For example: "'You suck and you should do this'...like, okay, is there a constructive way you can tell me that? Because you're really making me not want to play right now" (Participant 4) and "Our team captain...sometimes, and maybe I'm just completely reading this wrong, but sometimes I feel like he's a little...He's a little too aggressive sometimes" (Participant 13).

4.3 Understanding the Social Components of Learning Complex Games

4.3.1 The Hierarchical Nature of Co-Regulation

The most prominent takeaway from the results is the emphasis on how hierarchical relationships govern the CoRL process and how one's position in the hierarchy dictates the role one may play in others' learning process. Specifically, those higher on the hierarchy are permitted to co-regulate the learning process of those beneath them, but not the other way around. Since those perceived to be less skilled are lower on the hierarchy than those perceived to be more skilled, this means that less skilled players are often the targets of input from those who are more skilled.

As discussed earlier, the theory of co-regulated learning dictates how one who is more skilled may play an active role in a learner's learning process by handling the metacognitive elements of the task (i.e. helping them reflect on what they did or determine their next move) [273]. This system of having those with more experience guide those with less mirrors this phenomenon, suggesting that this is the most prominent way in which CoRL manifests among esports teams. This assertion is also supported by the fact that higher-skilled players may also dictate goals, a metacognitive process, for the less experienced, and help them monitor their progress toward those goals.

Interestingly, the results also suggest that players are particularly strict in upholding the hierarchy. Input is expected to come solely from those higher up and it is expected to be taken without much question by those lower down, with players going so far as to suggest that failure to accept this input can lead to conflict among team members. Further, when a player lower on the hierarchy attempts to give input to one higher, it is often seen as a social infraction, and may also result in conflict. Players seem to be consciously aware of these rules, as most participants suggested that they would never even try to give input to those perceived as better than them. Further, while CoRL suggests that this management from the more experienced is dialed back as the learner gains skill, our participants did not mention any such phenomenon, suggesting that the more skilled will always give input to those lower on the hierarchy in this way. It may be that the strict adherence to a hierarchy prevents players from perceiving those below them as having gained enough skill to not require their input. It may also be that players assume that even as they gain skill those above them continue to gain skill as well, thus remaining above them.

This strict adherence to a hierarchy is of further interest given that participants suggested that the hierarchy itself was not a constant and that there are contexts in which one's status may change. For example, a player who is, overall, lower skill level, may know how to play a specific character better than another player, giving them a higher skill level in that specific context, and permitting them to give input, but only in that context. This flexibility comes at an interesting juxtaposition with the strict adherence to a hierarchy as it raises questions regarding how players recognize such contexts and shifts in the hierarchical relationship. It also raises questions regarding how mismatched perceptions regarding the nature of the hierarchy at any given moment may lead to conflict. These questions can be explored further in future work, alongside questions about how this hierarchy manifests and is negotiated by players of different demographic groups. Furthermore, while players indicated their own strict adherence to this hierarchy, many had stories about *someone else* of a lower rank or status who had provided unsolicited advice, suggesting that not all players understand or adhere to the setup. This seems to indicate it may not be as strict as it appears.

4.3.2 The Relationship between Input and Failure

Participants also prominently discussed the relationship between input (sought or received) and recognition of mistakes or failures. Recognizing a failure or feelings that a mistake was made will prompt one to seek input while noticing repeated failure will prompt one to give input. Further, the input given tends to focus on identifying the specific mistakes, how they can be overcome, and how the player can do better next time. These findings mirror discussions in previous work about how difficult it can be for a learning player to recognize their mistakes and identify how to prevent them in future play [346, 349]. It similarly is reminiscent of work on CoRL in computing education, which saw a high level of social help-seeking behaviors, especially when the students felt they were hitting a wall [531].

Here we see that there is a strong social component to identifying and overcoming mistakes, something only briefly mentioned but not explored in previous work. This is also something that was prominently discussed as a key learning activity and a prominent challenge in the results of the previous study. Following Hadwin's theory, this demonstrates another way in which CoRL manifests among esports teams in which the more experienced can help manage a learner's metacognitive processes [273] since attribution (the identification of mistakes) and adaptation (how mistakes are overcome) are recognized as metacognitive elements of learning [756].

4.3.3 The Three Phases of Gameplay and Learning

Consistent with the results of the previous study, the findings here demonstrate that learning processes occur differently across the three phases of gameplay (pre-game, in-game, and post-game), but this time highlight the social component in more detail. Specifically, while pre-game is dedicated to discussing strategies and setting goals and post-game is dedicated to reflection and evaluation, no learning occurs in-game, with players instead focused on maintaining performance. While I have already discussed how these three gameplay phases mirror the three phases of learning defined by the Cyclical Phase Model of SRL [756] these results demonstrate that learning in esports cannot only be explained by individuals regulating their own learning but also through a complex set of interactions with other team members.

Hadwin's model proposes four phases: understanding, goal setting, working toward the goal, and adaptation. It appears that both phases one and two take place pre-game, while phase three is in-game and phase four is post-game. In phases one and two, the group negotiates their shared understanding of the task and their strategies for accomplishing it. In phase three, the group works collaboratively on their goal, collectively utilizing multiple cognitive, metacognitive, and motivational strategies. Finally, in phase four, the group makes small changes to large-scale pivots in strategies and goals based on feedback from the task and one another. Players in our study exhibited activity through all four of these CoRL phases via the role of group communication: setting goals and deciding on strategies through group discussion or respective hierarchies, motivational communication during gameplay, reflecting through group review of gameplay, and teammates holding each other accountable for setting and reaching goals.

The lack of in-game learning is interesting. It may be that esports players are not interested in learning in-the-moment, or it may be that learning is occurring and that they may not be aware of it. Here, our participants more explicitly described that there is no conversation around learning (i.e. evaluations or reflections) in-game, as they wish to focus all in-game communication on the gameplay itself and keep negative emotions as low as possible to avoid them impacting play. From this, I can, at least, infer that there is no *social* learning in game and that any subconscious learning that may occur happens at the individual level. This question can be explored further in future work.

4.4 Summary

While this is a mererly a first step towards expanding this work to understand social learning in complex gameplay, it illustrates how difficult it is for any player to learn in a vacuum and how much interaction with others, either directly or indirectly, matters when overcoming obstacles. The reality, however, is that many players are unable to go to others for aid. It is in these scenarios, then, that the presence of a computational tool that may support these players in the place of another becomes most beneficial. The results of this study provide initial insights into how a tool could be designed to support a player the way a real other person would and motivates the remainder of this work to focus, in more detail, on how solo learning occurs and interfaces with data-driven assistance.

Chapter 5

The Cyclical Phase Model in Complex Games

The work discussed in this chapter was originally published in Frontiers in Psychology in 2021 [1] [345]

5.1 Why the Cyclical Phase Model of Self-Regulated Learning?

Given the many models of SRL, I specifically chose, as I have already emphasized, to focus on Zimmerman's Cyclical Phase Model (CPM) [750] as the three-phase arrangement corresponds exceptionally well to gameplay. Specifically, the forethought phase, characterized by planning and goal setting, corresponds appropriately to any pre-game time-point that players may experience. For example, during esports play, players will, prior to beginning the match, set goals and plan strategies through the selection of specific characters, skills, or equipment. Alternatively, in a role-playing game (RPG), before engaging a boss monster or embarking on a quest, a player will choose a

[1]Kleinman, E., Gayle, C., & Seif El-Nasr, M. (2021). "Because I'm Bad at the Game!" A Microanalytic Study of Self Regulated Learning in League of Legends. Frontiers in Psychology, 12, 5570. This research was led by me but would not have been possible without the support of Christian or the guidance of my advisor Magy.

strategy and select party members, equipment, and items related to that strategy.

The performance phase, characterized by strategy execution and monitoring of progress, corresponds to the point in time at which players are "playing the game." In the case of the esport example, this would be while the player is in the match itself, during which time they will enact their chosen strategies and monitor their progress towards their goals. They may also, if able, adjust their strategy, such as changing what equipment is used or what character is played. In the case of the RPG example, this would correspond to the time during which the player is engaging the boss or pursuing the quest objectives. While doing so, the player would similarly execute strategies and monitor progress, adapting plans as needed. Notably, most games include interface elements, such as health or progress bars, that can help players monitor their progress.

Finally, the self-reflection phase, characterized by review and evaluation, corresponds to the point in time at which gameplay has "concluded". In the esports example, this would be after the match has ended, during which time most games present players with statistics regarding their performance. In the case of the RPG example, this would be after the completion of the boss fight or quest, at which point the player can evaluate the extent to which their chosen approach succeeded. In both cases, this phase will inform the next execution of self-reflection. In the case of the esport example, this would be the next time the player starts a match. In the case of the RPG example, this would either be the next boss fight or quest they encounter or, if the current task ended in failure, the next time they attempt it.

Based on this understanding, we can use CPM as a theoretical lens to understand learning, how it occurs, and how it differs across skill levels. Toward this end, I replicated an old study that investigated differences in the execution of CPM processes among athletes of different skill levels. Specifically, in 2002, Kitsantas and Zimmerman [343] used the Cyclical Phase Model to study differences in SRL between expert, non-expert, and novice volleyball players. They conducted a micro-analytic study in which players were asked questions about their general practice techniques for learning and mastering overhand serves. They were then asked to perform before the researchers and answer additional questions about how they felt they did and why they may have failed [343]. The results found that experts set better goals and had better planning during the forethought phase, better strategy use and self-monitoring during the performance phase, and better evaluations, attributions and adaptations during the self-reflection phase than either non-experts or novices [343]. I replicated this study in the context of League of Legends, replacing overhand serves with last hitting, and adapting the methods and measures to the context of the game, which I will describe in the following sub-sections.

5.2 Empirical Study of Self Regulated Learning Across Skill Levels: A Case Study of League of Legends

5.2.1 Methods

5.2.2 League of Legends

League of Legends is an esport game developed by Riot games and belonging to the Multiplayer Online Battle Arena (MOBA) genre. The game is played by two teams of five on a square map where each team has a base in either the lower left (for the green team) or upper right (for the red team) corner of the map. The bases house a crystal called a "Nexus" and the goal for each team is to reach and destroy the opposing team's nexus. The rest of the map consists of three lanes that extend from base to base and are referred to as top (for the one that follows the left and top edges of the square map), middle (for the one that cuts diagonally across the center of the square), and bottom (for the one that follows the bottom and right edges of the square map). There are also forested areas between the lanes, referred to as the "jungle". The three lanes each house six towers (three for each team) that fire lasers at opposing entities and must be destroyed in order to reach the enemy base. The jungle, by contrast, is home to various monsters that can be killed for gold or experience points.

In order to win a match of League of Legends, players must gain experience to level up their characters, gold to buy items to make their characters stronger, and win battles against enemy players in order to destroy the opposing towers and advance across the map. A key component of this process is a skill called "last hitting". Lanes in LoL are populated by small non-player entities called creeps that can be killed for gold and experience. If the player deals the finishing blow (as opposed to another player or allied creep doing so) they get *more* gold and experience. Intentionally attempting to deal these finishing blows is referred to as "last hitting".

5.2.2.1 Recruitment

30 League of Legends players were recruited to participate from collegiate esports teams, social media ads, and convenience sampling. The participants included 10 experts, 10 non-experts, and 10 novices, following the participant breakdown of the original study. Skill level was self-reported by prospective participants when filling out an online recruitment form.

5.2.2.2 Study Protocol

Participants joined a study session over zoom. They first provided demographic information, followed by a description of last-hitting used to assess their knowledge of the skill. All participants were then shown the same instructional video on how to execute a last-hit. They were then asked a set of questions regarding their self-efficacy, perceived instrumentality of last hitting, intrinsic interest in last hitting, goal setting, and planning. Following these questions, participants were instructed to open the League of Legends practice tool, where they could create a custom game and practice last hitting for ten minutes. Following the practice session, I asked participants about their strategy use, self-monitoring, self-evaluation, and self-satisfaction during the session. Participants were then tested for last hitting skill via a second game in the practice tool with the same arrangement. This time, however, they only last hit until they missed a last hit. All participants did miss a last hit. At this point they were asked about their attributions, adaptation processes, and self-efficacy perceptions.

5.2.2.3 Measures

The specific measures used and questions asked were as follows:

Last Hitting Skill: League of Legends tracks how many last hits a player has achieved in a user-interface (UI) element in the upper right corner of the screen. Last hitting skill was evaluated based on this number at the point at which the player missed the last hit during the second custom game.

Measures of Self Motivation: The questions for the measures of self-motivation were adapted directly from those used by [343]. All participants were asked the following questions to measure the respective factors:

- "On a scale from 0–100 with 10 being Not Sure, 40 being Somewhat Sure, 70 being Pretty Sure, and 100 being Very Sure, how sure are you that you are able to last hit every creep in a given wave?" (Self-Efficacy). This was asked once before practice and again after missing a last hit during the second custom game.
- "How interesting is last hitting to you on a scale from 0 to 100 with 10 being Not Interested, 40 being Somewhat Interested, 70 being Pretty Interested , and 100 being Very Interested" (Intrinsic Interest). This was asked once before practice.
- "How important is last hitting skill in attaining your future goals on a scale from 0 to 100 with 10 being Not Important, 40 being Somewhat Important, 70 being Pretty Important, and 100 being Very Important" (Perceived Instrumentality). This was asked once before practice.
- "On a scale from 0– 100 with 10 being Not Satisfied, 40 being Somewhat Satisfied, 70 being Pretty Satisfied, and 100 being Very Satisfied, how satisfied are you with your performance during this practice session?" (Self Satisfaction). This was asked once after practice.

Forethought Phase:

Goal Setting: Before practice, all participants were asked "Do you set any specific goals for your sessions when practicing last hitting and if yes, what are they?"

The researcher recorded the answer verbatim. The goals were then coded independently by two researchers into one of the following categories: outcome goals, technique of process goals, other, and no goals, the same scale used by [343]. For the context of League of Legends, the categories were considered as follows:

- "Outcome goals" referred to statements related to getting a certain number of last hits or amount of gold.
- "Process goals" referred to statements related to managing opponent presence or number and positioning of creeps in the lane.
- "Other" referred to any statements that did not discuss either of the above.

These definitions were developed and agreed upon by two researchers with extensive League of Legends experience. Cohen's kappa [133] was used to check for agreement and resulted in a score of .9, indicating very strong agreement [375].

Planning: Also before practice, participants were asked "Do you have a regular routine that you follow when you practice on your own?" The responses were again recorded verbatim and coded by two researchers into one of the following categories: completely structured routine, partially structured routine, or unstructured routine, the same scale used by [343]. For the context of League of Legends, these were defined as follows:

- A "completely structured routine" referred to discussions of regular practice using the practice tools or regularly playing warm up games in less competitive game modes.
- A "partially structured routine" referred to discussions of staying in practice by just playing regularly or irregular practice sessions.
- An "unstructured routine" referred to discussions of not practicing.

These definitions were developed and agreed upon by the same two researchers with extensive League of Legends experience. There were no disagreements in the code applications resulting in a kappa value of 1, indicating perfect agreement [375].

Performance Phase:

Strategy Use: Two questions were asked regarding strategy use, echoing Kitsantas and Zimmerman's protocol [343]. These were:

- "What do you need to do to accomplish your goals?" (Asked before practice)
- "What do you need to do to successfully execute the last hit next time?" (Asked after missing a last hit during the second custom game)

These were again recorded verbatim and coded by two researchers into one of the following categories: specific technique, visualization strategies, concentration strategies, both, and practice/no strategies, the scale used by [343]. For the context of League of Legends these were defined as follows:

- "Specific technique" referred to discussions such as getting the timing right, using the right skill, or targeting the right minion.
- "Visualization strategies" referred to any discussion of visualizing or imagining oneself doing it correctly.
- "Concentration strategies" referred to any discussion of focusing or concentrating either in general or on a specific aspect of gameplay.
- "Technique and concentration" referred to responses that included both.
- "Practice/no strategy" referred to answers that just discussed practicing or did not discuss any strategy.

These definitions were developed and agreed upon by the same two researchers. Cohen's kappa resulted in a score of .91 for the first question and .83 for the second, both indicating very strong agreement [375].

Self-Monitoring: After the practice session, all participants were asked "How did you monitor your performance and progress during the practice session?". These were again recorded verbatim and coded by two researchers into one of the following categories: creep score alone (corresponding to Kitsantas and Zimmerman's 'service outcome points alone'), use of technique or form and its outcomes, do not know, or other, the scale used by [343]. For the context of League of Legends these were defined as follows:

- "Creep score alone" referred to discussions of tracking the number of last hits achieved, either in one's head or using the UI's CS score board.
- "Use of technique or form and its outcomes" referred to discussions of technical execution of the skill, such as making sure the minions were in the right spot or managing their numbers.
- "Do not know" referred to statements indicating that they did not monitor their performance or were not sure if they did.
- "Other" referred to any self monitoring strategy that did not correspond with the above.

These definitions were developed and agreed upon by the same two researchers. There were no disagreements in the code applications resulting in a kappa value of 1, indicating perfect agreement [375].

Self-Reflection Phase:

Self-Evaluation: Also after the practice session, participants were asked "Did you evaluate your performance during the practice session? If so, how?" These were again recorded verbatim and coded by two researchers into one of the following categories:

- Self-evaluator (if they responded yes and gave a reasonable example of self-evaluation)

- Non-self-evaluator (if they responded no or failed to give a reasonable example of self-evaluation)

These are exactly the categories used by [343] and did not need to be adjusted to the context of League of Legends due to the general definitions. There were no disagreements in the code applications resulting in a kappa value of 1, indicating perfect agreement [375].

Attributions: After missing a last hit, participants were asked "Why do you think you missed the last hit?". These were again recorded verbatim and coded by two researchers into one of the following categories: form or technique, power, ability, practice, concentration, and do not know, the scale used by [343]. For the context of League of Legends, these were defined as follows:

- "Form or technique" referred to discussion of strategic failures such as wave or health management or player positioning.
- "Power" referred to discussion of physical failures such as reaction time or misclicks.
- "Ability" referred to discussion of one's gameplay skill.
- "Practice" referred to discussions of practice (i.e. needing more).
- "Concentration" referred to discussions of focus.

These definitions were developed and agreed upon by the same two researchers. Cohen's kappa resulted in a score of .78, indicating strong agreement [375].

Adaptation: After missing a last hit, all participants were asked the following three questions, answered with either a "yes" or "no", following Kitsantas and Zimmerman's protocol [343]:

- "After missing last hits, do you think about why you missed?"
- "When you miss a last hit, do you change anything during your next attempt?"

- “If you repeatedly miss last hits, do you ask your coach or teammates to give you feedback or advice?”

5.2.2.4 Data Analysis

Shapiro-Wilk tests were used to check for normal distributions of the numerical self-motivation data. Test results indicated that the data was not normally distributed, and thus non-parametric Mann-Whitney tests and Kruskal-Wallis tests were used for these data. Chi-square tests were used to assess differences for categorical data.

5.2.3 Results

5.2.3.1 Demographics

27 participants identified as male, 2 as female, and 1 as non-binary. Age ranged from 18 to 39. The average age for experts was 20.1, for non-experts was 21.9, and for novices was 24.5. On average, expert players had 5.75 years of experience, non-experts had 4.8, and novices had 4.9. Across the entire sample, 2 players played jungle (both experts), 8 played top lane (two experts, three non-experts, and three novices), 5 played mid lane (three non-experts and two novices), 9 played adc (five experts, one non-expert, and three novices), 5 played support (one expert, two non-experts, and two novices), and one (a non-expert) played fill (all positions).

5.2.3.2 Last Hitting Skill

Table 5.1: The means and standard deviations for creep score for each group.

Group	**Mean**	**STDEV**	**Median**
Experts	17	12.9	11.5
Non-Experts	15.1	19.1	9
Novices	8.4	12	5

Last hitting skill was determined using the number of creeps last hit by each player during the second custom game (when they were asked to last hit until they missed one). The means and standard deviations for each group are shown in Table 1. Kruskall-Wallis results indicate that the differences between all three groups are not statistically significant (all p > .05).

5.2.3.3 Measures of Self-Motivation

Table 5.2: The means and standard deviations for the five measures of self-motivation.

Variable	**Experts**	**Non-Experts**	**Novices**
Self-Efficacy (Before Practice)			
mean	79	58	58
stdev	14.5	21	15.5
Self-Efficacy (After Missing)			
mean	76	55	55
stdev	19	25.5	25.5
Intrinsic Interest			
mean	64	43	70
stdev	31	26.3	24.5
Perceived Instrumentality			
mean	91	79	82
stdev	14.5	20.2	25.3
Self-Satisfaction			
mean	73	73	58
stdev	17	17	25.3

The means and standard deviations for self-efficacy, intrinsic interest, perceived instrumentality, and self-satisfaction are shown in Table 2. Kruskall-Wallis results indicated that the differences between groups were not significant (P > .05) for all measures except for the Self-Efficacy (Before Practice) measure (H = 8.35, P = .01, Degrees of Freedom = 2). Mann-Whitney pair-wise test results with Bonferroni corrections indicate that experts had significantly higher self-efficacy at this point than novices (U = 79, P = .01). Non-experts did not differ significantly from novices or experts at this point (P > .016).

5.2.3.4 Forethought Phase

Goal Setting: There were significant differences in goal setting among the three expertise groups ($\chi^2(6) = 13.1$, P = .04). The counts for each goal type for each skill level can be seen in Table 5.3. Cramer's V was calculated to determine effect size and the result ($w = .46$) indicates a medium to large effect size.

Table 5.3: An overview of how different types of goals were set across the three skill levels.

Forethought: Goal Setting	**Experts**	**Non-Experts**	**Novices**
Outcome Goals	5	6	0
Process Goals	3	2	7
Other Goals	0	2	1
No Goals	2	0	2

Planning: There were significant differences in planning among the three expertise groups ($\chi^2(4) = 14$, P = .007). The counts for each goal type for each skill level can be seen in Table 5.4. Cramer's V was calculated to determine effect size and the result ($w = .48$) indicates a medium to large effect size.

Table 5.4: An overview of how different routines were used across the three skill levels.

Forethought: Planning	**Experts**	**Non-Experts**	**Novices**
Completely Structured	5	1	0
Partially Structured	4	5	2
Unstructured	1	4	8

5.2.3.5 Performance Phase

Strategy Use: There were no significant differences for strategy use before practice ($\chi^2(8) = 6.94$, P $> .05$) or after missing last hits ($\chi^2(4) = 4.26$, P $> .05$). The counts for each strategy type for each skill level can be seen in Table 5.5. For the second question, asked after missing last hits, "Visualization Strategies" and "Practice/No Strategy" were never applied to the participants' statements by the two researchers.

Table 5.5: An overview of how different strategies were used across the three skill levels at both question times.

Performance: Strategy Use	Experts	Non-Experts	Novices
Before Practice			
Specific Techniques	2	4	4
Visualization	2	0	0
Concentration	3	3	1
Technique and Concentration	1	1	1
Practice/None	2	2	4
After Missing			
Specific Techniques	5	6	6
Concentration	1	3	3
Technique and Concentration	4	1	1

Self-Monitoring: There were no significant differences between the groups for self-monitoring ($\chi^2(4) = 5.97$, $P > .05$). The counts for each technique for each skill level can be seen in Table 5.6. "Do Not Know" was never applied to the statements by the two researchers.

Table 5.6: An overview of how different self-monitoring techniques were used across the three skill levels.

Performance: Self-Monitoring	Experts	Non-Experts	Novices
Points	9	5	6
Technique	1	5	3
Other	0	0	1

5.2.3.6 Self-Reflection Phase

Self-Evaluation: There were no significant differences between the groups for self-evaluation ($\chi^2(2) = 2.2$, $P > .05$). The counts for self-evaluation for each skill level can be seen in Table 5.7.

Table 5.7: An overview of how self-evaluation occurred across the three skill levels.

Reflection: Self-Evaluation	Experts	Non-Experts	Novices
Yes	9	8	10
No	1	2	0

Attributions: There were no significant differences between the groups for attribution ($\chi^2(4) = 0.6$, P > .05). “Ability”, “Practice”, and “Do Not Know” were never applied to the attribution statements by the two researchers. The counts for the remaining attribution types across the skill levels can be seen in Table 5.8.

Table 5.8: An overview of attribution types across the three skill levels.

Reflection: Attributions	**Experts**	**Non-Experts**	**Novices**
Form and Technique	5	5	5
Power	3	3	4
Concentration	2	2	1

Adaptation: The responses for the three adaptation questions can be seen in Table 5.9. Chi square tests indicated no significant differences between groups (all P > .05).

Table 5.9: The number of people in each group who said yes and no for each of the adaptation questions.

Reflection: Adaptation	**Experts**	**Non-Experts**	**Novices**
Do you think about it?			
Yes	6	5	5
No	4	5	5
Do you change anything?			
Yes	4	7	7
No	6	3	3
Do you ask for help?			
Yes	3	4	4
No	7	6	6

5.3 The connection between Computational Support and Self Regulated Learning

In the study I just described, I observed that novices discussed process goals significantly more than non-experts and experts. This is in line with previous work that found that players seemed to shift from process to outcome goals as they obtained more skill [754]. That being said, expert players also placed a fair amount of emphasis on process when discussing their gameplay goals. For example “I’m not that focused on last hitting to get the minions because I find that somewhat easy, like it comes second nature to me now, there's other stuff I take into more account when I play and try to secure my farm. So like uh wave management, mainly, that's more important to me than last hitting to secure minions, and obviously just like not screwing up the lane and dying randomly" (Participant 16, Expert) and "[I will] see if I can get all of the CS when the wave is sitting in the middle, when I'm pushing, freezing, when I'm under tower. There's so many scenarios for where the minion wave is at and I want to make sure I can adjust and reach goals in every situation." (Participant 4, expert). This may suggest that, as players gain skill, they may start to shift back towards focusing on process over outcome. This may be because the desired outcome for last hitting is generally understood to be about 10 creeps per minute (for a total of 100 at 10 minutes). It may be that high-level players understand this as their desired outcome and revert to focusing on process in order to identify execution errors that may hinder it. While exploration of this phenomenon was beyond the scope of this work, it is still a trend worth noting.

Regarding planning, as a construct, Advanced players had significantly more structured practice routines than novice players. This is likely due to novice players being less likely to play on teams, which, as the previous study highlighted, have a notable impact on one's approach to gameplay and preparation for gameplay. This may also be because novice players are less interested in competitive play, and, as such, likely less interested in a formal or structured practice, instead choosing to play games for leisure. This is well articulated by Participant 22 (novice): "No, usually I just jump right into a game and go from there". These findings are consistent with those discussed by [343], suggesting that this is an area where esports and traditional sports share common ground. In other words, the results indicate that novice League of Legends players, like novice volleyball players, are more often engaged in casual play than structured training or learning.

There were, however, no significant differences in SRL processes for any other phase of SRL, which is in sharp contrast to the findings of [343], which indicated sig-

Figure 5.1: The League of Legends in-game UI presents information about player performance including kill counts, gold, experience earned, and creeps killed while playing the game.

nificant differences across all phases. A possible explanation for the overall similarity in SRL across skill levels in the performance and self-reflection phases may be found in the design of League of Legends itself. During play, the game tracks all participating players' progress information across a variety of metrics including gold amounts, level, enemy players killed, and, of course, creeps killed. This information is visible to any player in the game either in small menus on the border of the screen or through a dashboard that can be accessed with the press of a button, see Figure 5.1. Participant responses to the self-monitoring and self-evaluation questions indicated that they were using these interface elements to monitor their progress during play during the study and that they also do so on a regular basis when they play in their own environment. For example: "biggest thing is just looking at my CS vs. time elapsed" (Participant 19, non-expert) and "Mainly I just check the scoreboard, check my CS and stuff" (Participant 29, novice). Thus, it is possible that the design of the game itself is encouraging players to engage in self-monitoring practices whenever they play.

This may also be the reason for the lack of significant differences in the self-reflection phase. When a match of League of Legends ends, all players are brought to a post-game screen that depicts how much each player in the game contributed and offers an assessment of their performance based on a similar set of metrics as the in-game dashboard, see Figure 5.2. The game client also features a descriptive statistics interface that stores players' performance data over time and presents it to the player in aggregate graphs that depict, among other ways of evaluating, how the player performs in comparison to other players, see Figure 5.3. There also exist a number of third-party tools that present players with similar information, outside of the game client [601, 68, 459]. It is likely that the presence of such interfaces encourages players to reflect on their performance, especially the post-game screen, which is automatically shown to all players upon completion of a match. It is further possible that the community psyche encourages the use of such reflective practices when pursuing improvement. Players who are particularly motivated to improve at the game likely spend a fair amount of time interacting with these screens in order to extract actionable insights. In other words, these screens likely encourage players to engage in self-reflection processes. That being said, participant responses did not explicitly mention these screens or tools in the context of the study.

This assertion resonates with existing theoretical discussions on the role of visualized data within the player experience [280, 80, 449, 447] and personal informatics and quantified self in the context of games [361, 548]. Previous work has discussed how presenting players with data on their gameplay performance over time, motivates continuous play and facilitates improvement [80, 447]. The findings I present here suggest that the improvement that comes about as a result of interaction with this visualized data may be because the visualizations encourage the execution of self-regulated learning processes, although the players themselves may not be aware that they are leveraging such skills. This suggests further opportunities to support players seeking to improve gameplay through the development of visualizations of their gameplay data.

To summarize, based on the results I present here, I argue that players are engaging SRL processes in the performance and self-reflection phases because they have

Figure 5.2: The League of Legends post-game UI presents information regarding how each player performed during the game

Figure 5.3: The game client stores and aggregates statistical data to present players with overviews of their gameplay over time.

effectively been trained or influenced to do so through interaction with the game's interface or third-party tools. Because all players at all skill levels interact with the same interface, there are few significant differences. With this in mind, the significant differences in planning and goal setting (forethought phase) are likely the result of the game lacking any interface or interaction that supports SRL at that point in gameplay. This means that League of Legends itself provides little guidance on how to practice effectively, meaning that players must turn to external resources. Additionally, most of the third-party tools that exist for League of Legends do not aim to help players with goal setting or training routines. Existing literature acknowledges this phenomenon of players seeking out external resources [137, 661], but it may be that most novice players have not sought these resources, or not located the right resources, and therefore have not developed the same SRL skills for the forethought phase as their more skilled counterparts. Coaches ultimately emerge as the best resource for forethought phase skills, but non-expert and novice players are less likely to have access to coaches than expert players, resulting in significant differences in their knowledge and execution of these skills.

The differences between the results of this study and the results of [343]'s study are likely also influenced by the nature of the game examined and what it means to be a novice of that game. League of Legends requires players to complete a tutorial before beginning play, meaning that complete beginners, and certainly novices, possess some basic knowledge of gameplay and terminology. They also need to have taken the bare minimum initiative necessary to download the game and create an account. By contrast, volleyball novices recruited from a public court, such as those in [343]'s study, may or may not have ever looked at any formal documentation on how to play. While this does not necessarily make one game easier than the other, it does suggest an inherent difference in the knowledge level of novice players, which may explain why there was no statistical difference in knowledge of last hitting when there was one in knowledge of overhand serves.

5.4 Summary

Through this study, I ultimately propose that SRL occurs very differently in League of Legends, and likely in all esports and complex games that contain the relevant UI elements, compared to a traditional sport, due to the presence of computational support within the game's interface. Based on this understanding, I enforce the idea that computational support tools meant to support SRL can support players' learning and performance in complex games, making high-skill-level play and its benefits more accessible. This is the ultimate conclusion of this first thrust of my dissertation: an understanding of how learning works in complex games in terms of SRL, and how it is supported computationally. Based on this, I am able to move on to the second thrust, where I examine the existing state of the art of computational support from the perspective of SRL, and identify opportunities to improve it.

Part III

Supporting Self-Regulated Learning Skills in Complex Games through Computational Support

Following the work discussed in the previous part, I arrived at the conclusion that computational support tools, ranging from AI assistants to visualization systems, can, and already do, support SRL in complex gameplay environments, especially esports. Based on this conclusion, my next goal was to better understand how that support works and how it can be improved. In order to accomplish this goal, I embarked on the second thrust of my dissertation work, in which I mapped the state of the art of computational support for SRL in esports and explored gaps that, if addressed, could improve SRL support and, in turn, player performance.

Chapter 6

A Taxonomy of Intervention Types for Computational Assistants for Esports

The work discussed in this chapter and the next chapter was originally published at CHIPlay 2022 [1] [346]

6.1 Computational Support for Esports

Following the work described in the previous section, I came to the conclusion that gameplay interfaces that present players with data-driven, computational support often facilitate the execution of SRL skills in complex games, especially esports. This led to the derivation of RQ2: "How do computational tools support self-regulated learning skills in complex games?" However, existing literature had little discussion of this concept, as few had explored the intersection of computational support in learning. Instead, computational tools meant to support learning in games are often developed based on one of three high-level understandings of the gameplay experience: players require decision-making support, players analyze their gameplay afterward, and players

[1]Kleinman, E., Habibi, R., Yao, Y., Gayle, C., & Seif El-Nasr, M. (2022). " A Time and Phase for Everything"-Towards A Self-Regulated Learning Perspective on Computational Support for Esports. Proceedings of the ACM on Human-Computer Interaction, 6(CHI PLAY), 1-27. This research was led by me but would not have been possible without the assistance of Reza, Yichen, and Christian and the input and guidance of my advisor Magy.

use others' gameplay to learn [367, 11, 701, 319]. These assumptions inform the design of tools in the absence of a formal, empirical exploration of learning within the domain.

Regarding the first assumption, there exist a number of tools that attempt to provide statistically or algorithmically informed decision-support, typically in the form of recommendations or predictions [116, 115, 123, 184]. For example, Eger and Sauma [184] developed a system that leverages machine learning techniques to identify *Hearthstone* decks based on the first few cards they see, in order to help players quickly adjust their own gameplay based on their opponent's strategy. Their system examines the first few cards and provides players with a deck archetype that it believes the deck to be, effectively allowing players to strategize for what is to come [184]. In another example, Chen et al. [116] used Monte Carlo Tree Search to develop a recommendation system that could give players pre-game suggestions on who to play in multiplayer online battle arena games like *DotA2*. In many of these games, characters explicitly counter one another and choosing the right character or team of characters based on what your opponents have or could choose is a challenge. Through their system, Chen et al. relieve the cognitive load players face when attempting to make informed decisions in this context.

Regarding the second assumption, many tools leverage visualization towards helping players perform retrospective analyses of gameplay, such that they can identify mistakes and learn from their own gameplay or others' [701, 11, 319, 706, 708, 367]. One example of such a tool is Afonso et al.'s *Visualeague* [11]. This tool is designed to help *League of Legends* players review and reflect on gameplay. It does so through interactive maps and timelines that allow users to view how positioning, skill level, and item possession changes over the course of a match. By reviewing this information post-gameplay, players can develop a stronger understanding of what they did well or poorly during the game. Wallner and Kriglstein also emphasized the retrospective analysis of movement through spatio-temporal visualization in their user study on *World of Tanks* [701]. They developed and compared visualizations that showed each team on their own map with enemies only visible when they were within vision range, that combined the two teams into a single map, and that showed troop movement, major combat sites, and long-distance attacks using icons similar to what would be seen on a military map [701]. Their results found that readability, level of detail, and graphical design were most important to users and that the battle map visualization was rated most favorably overall [700, 701].

Regarding the third assumption, visualizations have also been leveraged towards helping spectators track the state of the game to support this process better [356, 113, 683, 351]. This work is based on an understanding that players can, and frequently do, learn by watching others play [625]. One prominent example of such work is that of Charleer et al. [112]. They designed spectator dashboards for *League of Legends* and *Counter-strike Global Offensive* based on spectator needs collected from a survey. To illustrate how their designs were informed by spectator needs: *League of Legends* spectators felt that gold was an important indicator of victory, and thus, the designed dashboard presented how much gold each player had at all times [112]. Their evaluations of their dashboards found that they improved spectators' experience of learning the game through spectatorship [112].

The primary drawback of all of the above work, however, is that, as discussed previously, computational tools for complex games are rarely designed or evaluated with learning, and especially CPM, in mind. Instead, much of the existing work focuses on usability [701, 706]. While the usability and design guidelines produced by this work are certainly valuable, we do not know if, or to what extent, the existing tools actually support learning. Learning, as we know, is a complex process composed of numerous low-level processes, including goal setting, planning, progress monitoring, and reflection [294]. Failure to support such processes at the right time may result in no support for learning or, in the worst case, could hinder it. For example, prompting reflection at the wrong moment could disrupt the overall progression of the learning process, break a player's focus in-game, and result in failure both for learning and gameplay. Further, having a comprehensive understanding of how existing tools support SRL is necessary to identify opportunities to enhance that process through the design of new or improved features.

I sought to address this gap through investigations of existing computational tools for gameplay learning. Again, I explored this question in the context of esports, for the same reasons discussed above, and because of the abundance of computational tools designed to support esports learning. Through this work, I aim to provide an overview of the ways in which existing tools support CPM processes, players' preferences and needs for support in the context of CPM, and opportunities to more effectively support learning in the context of CPM.

6.2 A Systematic Review of the State of the Art

My initial goal in this thrust was to provide an overview of the state of the art of existing tools from the perspective of CPM. Thus, I first conducted a systematic review of existing computational assistants for esports in order to develop a taxonomy of intervention types. Specifically, this review sought to answer the following question:

- What kinds of interventions do existing assistants for esports offer to support learning during each phase of learning according to the Cyclical Phase Model of Self-Regulated Learning?

This review occurred in two parts. In the first part, commercially available, third-party tools were reviewed to define and validate the types of interventions that they offered. In the second part, the tools were re-reviewed to determine when, before, during, or after gameplay, each tool offered each intervention type.

6.2.1 Methods

To be included in the analysis, the tool had to be commercially available, free to use, and offer support beyond just stat tracking. These criteria were selected to capture the state of the art of computational support for esports as it is experienced by the majority of players. Tools that are not commercially available, such as those exclusive to the research literature, are more difficult, if not impossible, to acquire and access. Similarly, while many paid tools, or the subscription variants of otherwise free-to-use tools, offer more features, many esports players may be unable or unwilling to pay. Further, while many tools exist that simply track and display stats over time, I was interested in tools that offered more advanced or explicit assistance. I additionally excluded any tools or systems that exist within the game client itself. These tools, especially those interacted with during gameplay, face different design considerations and restraints than the development of external tools, as they need to interface with the game and its mechanics. For the purposes of this work, I thus focused on third-party tools that the player interacts with outside of the game itself.

Tools were identified through a search procedure leveraging combinations of the following keywords: "computational assistant(s)", "assistant(s)", "AI assistant(s)" and "esports". Searches also combined the aforementioned keywords with the titles of popular esports games. To determine which game titles should be included, I searched Google for "top 10 esport games" and included the titles that appeared on more than one list on the first page of the search results. Once a set of keywords was inserted, I opened the web page for every assistant that appeared on the first page of the results. These were checked to ensure they met the aforementioned inclusion criteria and added to the sample if they did. Tool identification stopped when the searches produced no new tools.

Based on these inclusion criteria and search methods, I identified seven tools: Senpai [601], Mobalytics [459], Fridai [228], Blitz [68], GOSU [256], OP Desktop App [489], and Porofessor Desktop App [525].

I installed and used each tool, in the context of *League of Legends*, taking note of all features, and referencing online resources to note any features not apparent through my own use. Through an iterative process similar to thematic analysis, I combined and collapsed these features into nine distinct intervention types. From this initial analysis, conducted with the versions that were public for these tools in May 2021, I generated a nine (intervention types) by seven (tools) matrix identifying the features offered by each tool. For validity, a collaborator generated another matrix independently, which was then compared against mine. There were 9 disagreements out of 63 marks, indicating a percent agreement of 86%. The disagreements were all resolved through a discussion.

I then added the relevant time points, before, during, or after play, to the matrix. Following the same process as above, two collaborators developed two matrices independently, and I then compared them to measure the level of agreement. There were 27 disagreements out of 216 total marks (3 time points x 9 intervention types x 7 tools), indicating a percent agreement of 88%. I led iterative discussions with both collaborators to resolve disagreements and generate the final matrix.

6.2.2 Results

I identified nine intervention types shown in Table 6.1. Table 6.3 summarizes the findings by showing how many (with a qualitative value of most, all, few, or none) tools offer each intervention during each CPM phase.

Intervention	***Definition***
Predictions	Offer a guess as to what is to come
Recommendations	Offer a suggestion as to what you should do
Updates	Inform the player when something new has happened
Reminders	Persistently display game information or state details
Reflections	Present an overview of gameplay and performance stats right after the match
Retrospections	Aggregate gameplay stats from multiple past matches
Instructions	Present step-by-step information on how to complete tasks
Sharing	Facilitates easy sharing of data or gameplay in some way
Evaluations	Present a judgement of performance quality

Table 6.1: The nine intervention types, their definitions, and examples.

		Pred.	***Rec.***	***Upd.***	***Rem.***	***Ref.***	***Retro.***	***Inst.***	***Shar.***	***Eval.***
Blitz	*Forethought (Before)*	✓	✓							✓
	Performance (During)	✓	✓		✓			✓		
	Self-Reflection (After)					✓	✓		✓	✓
Fridai	*Forethought (Before)*									
	Performance (During)		✓	✓	✓			✓	✓	
	Self-Reflection (After)									
GOSU	*Forethought (Before)*		✓							
	Performance (During)	✓	✓	✓	✓			✓		
	Self-Reflection (After)		✓			✓			✓	✓
Mobalytics	*Forethought (Before)*	✓	✓		✓		✓			✓
	Performance (During)	✓	✓	✓	✓					✓
	Self-Reflection (After)		✓			✓	✓		✓	✓
OP	*Forethought (Before)*	✓	✓					✓		
	Performance (During)		✓		✓					
	Self-Reflection (After)					✓	✓		✓	
Porofessor	*Forethought (Before)*		✓				✓			
	Performance (During)	✓	✓	✓	✓		✓			✓
	Self-Reflection (After)					✓			✓	✓
Senpai	*Forethought (Before)*	✓	✓				✓			✓
	Performance (During)	✓	✓		✓					✓
	Self-Reflection (After)		✓			✓	✓		✓	✓

Table 6.2: Distribution of interventions per phase for the reviewed tools.

	Pred.	***Rec.***	***Upd.***	***Rem.***	***Ref.***	***Retro.***	***Inst.***	***Shar.***	***Eval.***
F	Most	Most	None	Few	None	Few	Few	None	Few
P	Most	All	Most	All	None	Few	Few	None	Few
R	None	Few	None	None	Most	Most	None	Most	Most

Table 6.3: How many of the reviewed tools (all, most (more than half), few (less than half), or none) offered each intervention during each phase (F=Forethought, P=Performance, R=Self-Reflection.

6.3 The State of the Art of Computational Support for Esports

In this section, I describe the state of the art based on the results of the study, which highlighted how each of the interventions in the taxonomy is implemented in existing tools.

6.3.0.1 Predictions:

During the forethought phase, most tools featured predictions of what positions or characters enemy players will select and of the chance of a character on the player's team winning against another character on the enemy team. Therefore, these interventions inform planning, e.g., if a player sees that character A is unlikely to beat character B, they can pick character C instead.

Most of the tools offered predictions during performance that showed the chances of a given character beating an opposing character. Notably, however, these were primarily static, and do not update as the game plays out. The only tool that offered dynamic predictions was GOSU [256], which offered in-game predictions of what the opponents may do next, such as ganking mid-lane. Notably, GOSU has shut down since the time of this analysis.

None of the tools offered predictions during self-reflection.

6.3.0.2 Recommendations:

During the forethought phase, most tools offered recommendations consisting of either bans, specifically which characters should be rendered unplayable, or load-outs, including characters, stats, and abilities. In most cases, the recommendations at this point were based on community trends.

All tools offered recommendations during the performance phase. In most cases, these were recommendations to level certain skills or build certain items. In some cases, tools recommended when to perform a certain strategic maneuver, such as placing a ward. Notably, however, like predictions, recommendations were primarily static.

Only a few tools offered recommendations during the self-reflection phase, which were mostly connected to evaluations, and would couple the identification of a player's weaknesses with suggestions for how to address them in future play.

6.3.0.3 Updates:

No tools offered updates during forethought or reflection. Most offered them during the performance phase, typically in the form of notifications, on the player's screen or via audio, that a certain objective had been met. For example, Mobalytics [459] informed players when any given player in a game of LoL had reached level 6.

6.3.0.4 Reminders:

Mobalytics was the only tool that offered reminders during the forethought phase. It tracked and made the player aware of which characters were banned, chosen, or still available.

All of the reviewed tools offered reminders during the performance phase. Here, reminders tracked the status of the players in the game (e.g., their health, level, and resources), and the status of NPCs. They displayed this information to the player either via an overlay or another, typically an internet browser, window.

No tools offered reminders during self-reflection.

6.3.0.5 Reflections:

Unsurprisingly, reflections were only offered during the self-reflection phase, with only Fridai not offering them. Reflections typically presented aggregated stats for all players on both teams. These included basic kill and death counts, the amount of damage done, resources obtained, objectives met, and any kinds of unique achievements.

6.3.0.6 Retrospections:

A few tools offered this information during the forethought phase, where players would be presented with their win rates in certain roles, on certain maps, or with or against certain characters. Typically this came in the form of a single percentage value, and, in some cases, players could see more detailed information about their past gameplay on a different window.

One tool, Porofessor, offered retrospections during the performance phase. Similar to the forethought phase, it showed players a single indicator of their past win rate with or against a given character.

Unsurprisingly, most tools offered retrospections during the self-reflection phase, giving players the ability to compare recent and past gameplay. Retrospections in this phase typically offered more detail, such as charts and graphs depicting statistical information such as resources obtained or the number of kills and deaths with a given character across past matches.

6.3.0.7 Instructions:

The OP.GG desktop app was the only reviewed tool to offer instructions during the forethought phase. Specifically, when a player selected a certain character, the tool would present step-by-step instructions for how to use that character's skills to execute strategic maneuvers.

A few tools offered instructions during the performance phase, where they offered step-by-step guidance for the execution of maneuvers. They functioned like an instruction manual.

None of the reviewed tools offered instructions during self-reflection.

6.3.0.8 Sharing:

Sharing was only offered during the self-reflection phase. In almost all cases, sharing occurred in an indirect manner, by uploading a player's gameplay data for a recent match to either a public database or by giving them the means to share it on social media. Through this, players were provided with a way to show their gameplay to others.

6.3.0.9 Evaluations:

A few tools offered evaluations during the forethought phase. These typically were present in the form of a brief, one-sentence description of the player's known strengths and weaknesses based on their past gameplay, often the game they had just played most recently.

Similarly, a few tools presented evaluations during the performance phase, Typically, these were alike to those presented during the forethought phase. During both the forethought and performance phases, there were a couple of tools that offered evaluations for other players, including opposing players.

Evaluations were offered during self-reflection by most tools, in the form of a judgment or grade. Typically, these evaluations were focused on finding the player's notable weaknesses in the just-played match, for example, poor vision. Often, these would be coupled with a recommendation for addressing these weaknesses.

6.4 Summary

To my knowledge, this work was the first to map the state of the art of computational support for esports and the first to examine this support through an SRL lens. In the context of this dissertation work, this taxonomy provides us with a means by which the features of computational support tools can be described and the tools themselves can be compared against one another. In doing so, we are better able to understand the range of support offered by existing tools and, once we collect the information, better able to compare it against players' needs and identify areas for improvement, which will be the topic of the next chapter.

Chapter 7

Deriving Design Requirements for Computational Support for Esports

The work discussed in this chapter and the previous chapter was originally published at CHIPlay 2022 [1] [346]

7.1 The Need for a User-Centric Approach to Computational Tool Design for SRL

While a great deal of work exists that examines or proposes new systems for computational support for complex games and esports, very little work takes a truly user-centric approach. In the cases of some tools, players, the intended end-users of the systems, play no role in the development process. This is often the case for purely algorithmic systems, such as SENSEI [319], described as an intelligent advisory system intended to help esports players improve their gameplay through advanced analytics and ML. There is also the work of Christiansen et al. [123] who developed a novel approach for measuring the causal effects of game features or player performance on chances of

[1] Kleinman, E., Habibi, R., Yao, Y., Gayle, C., & Seif El-Nasr, M. (2022). " A Time and Phase for Everything"-Towards A Self-Regulated Learning Perspective on Computational Support for Esports. Proceedings of the ACM on Human-Computer Interaction, 6(CHI PLAY), 1-27. This research was led by me but would not have been possible without the assistance of Reza, Yichen, and Christian and the input and guidance of my advisor Magy.

winning, motivated by a desire to help players and developers better understand the connection between a given set of statistics and victory. This also applies in the case of the work of Chen et al. [116, 115], discussed previously. All of these examples involve systems designed, specifically, to help players improve at play, either by tracking their performance, helping the understand the outcomes of their choices, or recommending the best approach. However, none of these works evaluate their systems with players or perform any explicit design requirements gathering.

Systems that include visualization tend to be developed in slightly more user-centric manners. Van den Broek et al. [683] explored the value of including tangible elements in a data visualization system for esports, and found that most potential users saw value in the approach. Additionally, Wallner et al. [703, 704, 707, 701, 706], Kriglstein et al. [366], afonso et al. [11, 10] and Halabi et al. [274], who all developed a number of different visualizations for gameplay data [703, 707, 274, 701, 366, 706], have evaluated their various systems with users as a part of the design and development processes.

There are even those works that did seek to develop more explicit design requirements, to better understand what players need. For example, Kuan et al. [367] began the development of their visualization system for *Starcraft* data with a requirements analysis with experienced players. They determined several user needs that their system had to address (such as identifying important events) and were able to use these to inform the design of the system. Their requirements, however, are specific to the context of *Starcraft* and may not generalize to other games. On a more general note, Wallner et al. [708] conducted a large-scale interview and survey study with players of first-person shooter, MOBA, and strategy games to better identify players' wants and needs from post-play game data visualization systems. These results generalize better across game genres but are specific to the use case of post-game data analysis.

In contrast to the gaming literature, the benefits of taking a more rigorous, user-centric approach to learner-facing computational support tool design are well documented in the domain of learning analytics dashboards, where tools are often designed to support students in classroom environments. For example, before creating "StepUp!", Santos et al. [578, 579] conducted a requirements analysis study with students enrolled in three university-level engineering courses. This study comprised of three brainstorming sessions, one with the students enrolled in each course, during which students were asked to identify issues they experienced. They were then asked to rate the issues in terms of how important it was to address them through a LAD. These sessions resulted in 34 issues. From these, the authors selected those that were rated the highest and, more specifically, that could be addressed by the StepUp! System (i.e. lack of a good learning environment was not something they could address). Issues they did select include such decision-making elements as "how I distribute my time" and "being aware of which resources and tools everyone is using".

Park et al. [501] followed a similar method for their requirements analysis but conducted one on one interviews instead of group brainstorming sessions. Their results found that students wanted to be able to see information regarding their learning patterns as this would help them make decisions around planning their learning schedule, managing their learning process, and setting their learning goals. Further, they wanted to be able to see accurate and trustworthy information, which they acknowledged might be more accurate than what they themselves perceived. Additionally, students were not keen on a tool that would analyze their learning pattern or disclose it to their teacher, indicating that, while they were open to using the tool to inform decision-making, they felt an aversion to evaluation. They were also in favor of seeing statistics about their scores and liked the idea of comparing themselves with other students.

Both StepUp! and LAPA were designed to meet these needs and support student decision-making [578, 579, 501]. Here, I argue that advancing the state of the art of computational support for learning in complex gameplay, especially esports, requires a similar, user-centric approach to requirements analysis. Currently, there is no concrete understanding of players' requirements for computational support for SRL in this context. Without such an understanding, it is difficult if not impossible to identify opportunities to improve existing design conventions or evaluate the extent to which the tools meet players' needs. Therefore, in this chapter, I detail a study I conducted that sought to derive user requirements for the design of computational support tools for SRL. By comparing the resulting requirements to the results of the taxonomy, I hoped to identify concrete opportunities to improve the state of the art.

7.2 A Mixed-Methods Examination of Players' Support Needs

To address this point, I built on the work discussed in the previous chapter and conducted a mixed methods study that sought to answer the following question:

- What kinds of interventions do players desire during each phase of learning according to the Cyclical Phase Model of Self-Regulated Learning?

The ultimate goal of this work, as stated above, was to gather user requirements for computational support systems meant to support learning in complex games, specifically esports, and to do so through the lens of CPM. In order to reach this goal, I first distributed an online survey to collect players' preferences in terms of when, before, during, or after gameplay, they wished to interact with each of the intervention types identified in the taxonomy discussed in the previous chapter. As with the systematic review discussed above, here, we understand that before gameplay corresponds to forethought, during to performance, and after to self-reflection. I followed this survey with data-driven retrospective interviews to gain deeper insights into how players leveraged computational assistance to support SRL, how they understood such assistance did or could help them, and what their preferences and needs were in this context.

7.2.1 Methods

7.2.1.1 Survey Design

The online survey first collected demographic data, then asked players about their preferences regarding each intervention type. For each intervention type, the participant was first provided with the definition as such:

- **Predictions:** Predictions offer a guess as to what is to come. For example, a tool may predict your chances of winning, the odds that a given character will be played in a given position, or the odds that the opponents are about to execute a certain move.

- **Recommendations:** Recommendations offer a suggestion as to what you should do. For example, a tool may recommend that you play a certain character or position, build a certain item, execute a certain move during play, or practice certain skills in future games.

- **Updates:** Updates inform the player when something new has happened or a milestone has been reached. For example, an update may inform you that an enemy player has reached a certain level.

- **Reminders:** Reminders persistently present game information or state details so that players can be aware of them. For example, a reminder may tell you which characters have been chosen and what their win rates or positions are, how many cards have been removed from a deck, or your own score as a player.

- **Reflections:** Reflections present an overview of gameplay and performance stats in a recently played match, typically right after the match completes. For example, a reflection may present an overview of how much damage each player did in the match.

- **Retrospections:** Retrospections present past gameplay information to prompt players to track or think about their performance in older games. For example, a tool may track gameplay over time and present an overview of how well a player performs in a given role or as a given character.

- **Instructions:** Instructions present what to do and how to do it. For example, a tool may provide a player with the locations of items that need to be acquired to complete a quest or instructions on what is needed to build an item, or the step-by-step maneuvers needed to execute a move.

- **Sharing:** Sharing specifically refers to whether or not the AI assistance facilitates easy sharing of data or gameplay in some form. For example, an AI assistant may automatically upload data to a database where other players can see it or may make it easy to record gameplay to share with others.
- **Evaluations:** Evaluations present a judgment of performance in some way and are meant to tell a player where they performed well or poorly. For example, a tool may tell a player that they were under-warding or did not get enough gold.

They were then asked the following question:

- When, during the gameplay experience, have, or would, you use [intervention type]? Circle all that apply (Before, During, and After Gameplay)

7.2.1.2 Survey Distribution

The survey took 15 - 20 minutes to complete. Participants were compensated 20$ for their time. The survey was distributed to the Senpai.GG community via a banner advertisement embedded in the application and was also shared via social media and word of mouth. Respondents from Senpai.GG were also able to share the survey with friends who were not a part of the Senpai.GG user-base. Participants needed to be 12 years or older, with those under 17 asked to provide parental consent, to participate. Though we hoped to recruit as many participants who had experience with computational tools as possible, this was not a requirement to participate. Participants were allowed to be located anywhere in the world, but the survey was only offered in English.

7.2.1.3 Interview Protocol

To augment the survey results, I conducted qualitative interviews with survey respondents. I recruited seven participants via email from the survey participant pool. Only respondents 18 years of age or older could participate. Interviews were conducted using Data-Driven Retrospective Interviewing [186], a method in which the participant is shown data about their own or others' actions or performance and asked to elaborate on what they think it means or how they interpret it.

During the interview, each participant was first shown their own survey responses compiled in a Google Sheets spreadsheet and asked:

- "Can you elaborate on why you prefer these interventions during the forethought phase and how you would use them?"
- "Can you elaborate on why you prefer these interventions during the performance phase and how you would use them?"
- "Can you elaborate on why you prefer these interventions during the self-reflection phase and how you would use them?"

The participant was then shown the data seen in Table 7.5 and asked:

- "Can you say a bit about why you think players prefer these interventions during the forethought phase and how you, or players like you, would use them?"
- "Can you say a bit about why you think players prefer these interventions during the performance phase and how you, or players like you, would use them?"
- "Can you say a bit about why you think players prefer these interventions during the self-reflection phase and how you, or players like you, would use them?"

7.2.1.4 Interview Analysis

Similar to earlier studies, the interviews were transcribed using a text editor. The data was segmented into lines based on how the text lined up in the editor with lines with five or fewer words combined with the previous line. Each line was treated as a unit of analysis. Interview data were analyzed using iterative thematic analysis and line-by-line coding [237, 574]. I performed open coding [574] on a representative sample consisting of 30% of the data, defining a set of themes regarding participants' use of each intervention type during each phase. I then worked with a collaborator to, separately, apply the themes to a different 30% of the data set in order to measure validity through Cohen's Kappa [133]. The resulting IRR was .74, indicating strong agreement [375]. I then coded the entire dataset.

7.2.2 Survey Results

The survey received 116 complete responses. An overview of the demographics can be seen in Table 7.1. Gender, race, and nationality, were not collected in order to prevent biasing of the results and to mitigate the risk of identification. The games respondents played can be seen in Table 7.2 and their experience with the seven reviewed tools can be seen in Table 7.3. Survey respondent preferences for which interventions they wanted during each phase can be seen in Table 7.5. Additionally, the games played by the seven interview participants can be seen in Table 7.4.

Age	**Under 20**	**20-30**	**Over 30**
#	62	46	8
Years of Experience	**¡ 5**	**5 - 10**	**¿ 10**
#	80	32	4
Expertise (self-reported)	**Novice**	**Intermediate**	**Expert**
#	19	76	21

Table 7.1: An overview of demographic data collected by the survey.

Game:	League of Legends	World of Warcraft	Dota2	Valorant	Counterstrike	Auto Chess	Various Fighting Games	Starcraft	Overwatch	Rocket League	Clash Royale	Fifa
#:	80	1	2	9	12	1	1	1	2	3	3	1

Table 7.2: The games played by the survey respondents and the number of respondents per game.

Tool:	SenpAI [601]	Mobalytics [459]	Blitz [68]	OP [489]	Fridai [228]	Gosu [256]	Porofessor [525]
#:	37	3	16	1	3	15	0

Table 7.3: The number of survey respondents (out of the total 116) who had experience with the reviewed tools.

Participant	1	2	3	4	5	6	7
Game	Starcraft	Dota2	Auto Chess	LoL	LoL	Fighting Games	LoL

Table 7.4: The games played by the seven interview participants.

	Pred.	***Rec.***	***Upd.***	***Rem.***	***Ref.***	***Retro.***	***Inst.***	***Shar.***	***Eval.***
F	60%	59%	34%	43%	28%	37%	39%	27%	28%
P	45%	50%	37%	47%	28%	34%	47%	24%	28%
R	53%	52%	35%	38%	67%	61%	32%	45%	59%

Table 7.5: An overview of what percentage of players preferred a given intervention during a given phase, according to the survey responses. F=Forethought, P=Performance, R=Self-Reflection.

7.2.3 Interview Results

The following results dictate, specifically, what users want in terms of computational support in terms of each intervention identified in the previous chapter during each phase of CPM. In other words, these results present a thorough set of design requirements that can help inform and guide the development of new and innovative tools.

Predictions: Participants indicated that, during forethought, they specifically preferred predictions regarding their opponents' strategies, so they could start planning their response. For example, Participant 6 stated: "[predictions] would come more as what you should expect from a certain kind of character and a certain kind of archetype that you're playing against. So in those terms, you're going to start predicting what a character, what an opponent, is going to do to you, and planning how to respond."

Participants wanted help predicting the actions of an opponent during the performance phase so that they could adapt their strategy to respond earlier. For example, Participant 5 stated: "If I knew what towers, for example, would be destroyed or aimed at or if I knew that the enemy players would be aiming at that tower I would either have our team protect it or go take another tower."

Participants indicated that there were two ways that predictions could aid them during self-reflection. The first was when predictions could give players an idea of what the opponents might do in the next match. Participant 4 explained: "predictions after make sense if you play a couple games in a row, Valorant or CS GO play that way, where you play best of five, so maybe what weapons they're going to buy or which station they're going to go to." The second was seeing predictions while reviewing a replay of the match, which could allow them to see how their odds of winning changed as they made decisions. This is discussed by Participant 2: "if I was getting predictions on the match while I was reviewing the match, and seeing throughout how the match may swing based on the events that occurred, that could be useful."

Recommendations: Participants indicated that recommendations could help them choose a specific plan or strategy. Participant 4 discusses this in terms of LoL's character selection phase, stating: "recommendations would be, for example, to counter-pick someone. Especially if I don't know the character the enemy picked, to be able to play against them."

Participants indicated that recommendations, during performance, would provide similar aid to that which they provided during the forethought phase. For example, Participant 7 stated: "It would be nice, during the game, just a recommendation to the player, specifically when they're new, "hey, look, you're playing against 17% magical damage, so stop buying armor" for example."

Participants indicated that, during self-reflection, recommendations would be connected to evaluations of their gameplay. For example, Participant 6 explained: "an AI that can evaluate that and see how many times you've fallen for that and be like "you can totally just block here and do a light quick punch and block this combo stream". That'd be amazing..."

Updates: Updates were preferred by 34% of respondents during the forethought phase. Unfortunately, however, no interview participant discussed updates during the forethought phase.

Participants indicated that updates, especially updates regarding the status of opposing players, could help them adapt their plans during performance. For example, Participant 3 said: "so like "x player just hit their big whatever alert alert" and that's helpful because it's hard to track seven people's boards at once."

35% of respondents preferred updates during the self-reflection phase. However, interview participants did not discuss updates during this phase.

Reminders: Participants indicated that forethought reminders could help them keep track of information, and reduce the cognitive load needed to do so, while making plans during this phase. For example, Participant 2 said "to maybe be reminded that certain characters are still in the selection pool or whatnot, just so you don't pick something incorrectly, forgetting about that one other character that might counteract your whole strategy."

Participants similarly discussed performance-phase reminders in terms of keeping track of game state information and reducing cognitive. For example, Participant 4 stated: "for example, if you would play hearthstone...sometimes you don't remember all the cards you have, so it's hard to keep track of what you can still use and what is out of the game."

38% of respondents preferred reminders during the self-reflection phase. Unfortunately, however, none of the interview participants discussed reminders during this phase.

Reflections: Participants acknowledged that being able to review a recent game during forethought could be helpful in deriving plans for a new game, as it could help them identify gameplay habits they should be attentive of. For example, Participant 5 said: "they could see what they did bad or good in the previous game if they're playing the same champion in the next game."

28% of respondents preferred reflections during the performance phase. While not discussed extensively by interview participants, Participant 5 did suggest in-game reflections related to a specific moment: "you don't really need them during the game, unless you want to get reflections or retrospections right away after you either die or make a mistake."

Participants discussed how reflections during the self-reflection phase were critical to their ability to build causal relationships between what they did and the ultimate outcome of the game. For example, Participant 7 said: "After the game...you have to know if you helped your team, if you healed your team at the right time, how much, your presence with your team or if you're a split pusher if you helped your team get the objectives, if you did things alone because you're a tank and you can do it yourself."

Retrospections: Participants saw potential for players, during forethought, to be able to use retrospections to inform their planning. For example, Participant 1 said: "So if you're playing Protoss it'll be like "you win 50% of the time against Zerg" or something and I think it's useful because you can be like "oh I'm fighting against this it'll be an uphill match".”

Only one interview participant, Participant 5, discussed using retrospections during the performance phase, discussing how retrospections could make them aware of their past mistakes during gameplay: “retrospections during the game, might remind us of our past mistakes or something like, if we die because of a mistake we did, and we did that also in the past, it could remind us that we had bad position or did little damage because we were in the middle of the fight, or because we didn't use our ult in time.”

Participants discussed how retrospections during self-reflection could be used to track progress towards improvement or identify negative habits. Participant 1 discussed this, stating “a tool that shows you how fast you expand how fast you get your next base just seeing how...theoretically if you're getting better and you want to take that next base earlier you'd want to see that time-point go down so being able to see that would be useful.”

Instructions: Only Participant 6 discussed instructions during the forethought phase, where they described how a system could provide the player with instructions outlining what a chosen character could do: “instructions that would basically be learning what you're inputs are and what you're character is capable of.” Participant 6 did not, however, expand on how they would use these instructions.

Participants discussed performance phase instructions in terms of plan execution. For example, Participant 2 said: “especially for newer players, taking on a game or a new position in a game you may not know it very well, having instructions to guide you on basics could be very useful to a player, especially in training.”

Only Participant 6 discussed instructions during the self-reflection phase, where they discussed them being offered in tandem with recommendations: “you're reflecting on what happened and you saw that the player was doing a specific thing and you didn't know what to do during the match. Yeah, that's the moment when you start getting instructions and recommendations.”

Sharing: Only Participant 5 discussed sharing during the forethought phase, in terms of obtaining information from others, either teammates or opponents, that could inform planning: “I get stats and also items, what the other team built, from another game, and how many kills, deaths, and assists they got in the last game...”

24% of respondents preferred sharing during the performance phase. Unfortunately, none of the interview participants discussed sharing during this phase.

Participants discussed the benefits of sharing during self-reflection as a way to gain deeper insights into their gameplay through communal review and reflection. For example, Participant 4 stated: “if there was a misplay, you and your ADC, for example, can look over the data and be like “why did you go there?” and I would like that because I have this crazy ADC who goes in, and I'm slow, and he dies, and then I'm sad. So that would be really good to improve as a team rather than as just a single player.”

Evaluations: Participants indicated that having pre-game evaluations, based on past gameplay, could help with planning for the upcoming match. For example, Participant 1 said: “Like you could evaluate and be like “oh maybe you've got Protoss vs. Zerg” and “the Zerg is probably going to rush you” and then the evaluation could be like “you suck at responding to rush” and then you'd be like “ok I better be ready”.”

Participant 5 discussed how in-game evaluations could provide information on a single moment, immediately afterwards: “after you die, after making a mistake, it could be good to get an evaluation of what you could have done better.”

Participants indicated that self-reflection phase evaluations were beneficial as a source of guidance, for example, Participant 5 said “without them (evaluations), I would have to find those mistakes on my own, but the thing is for a long time I got stuck with a champion in mastery five because I didn't know what I did wrong...My mistake was that I didn't get a lot of farm. And that I also didn't ward a lot. That was what helped. It would've helped if I'd known that before.”

7.3 Opportunities to Better Support SRL

Forethought		
What Players Do/Want Help With During this Phase	***What Existing Tools Offer***	***Implications and Opportunities for Future Research and Design***
Determining Opponents' Strategies, How to Respond, Tracking Information, Making Informed Decisions while not Dwelling on Past Mistakes	Predictions of Load-out or Victory, Recommendations and Instructions for Executing Strategies, Information Tracking, Overviews of Past Performance	Detailed Strategic Predictions, Goal Oriented Support, De-Emphasized Review

Table 7.6: A summary of what players do and want help with in the Forethought phase of SRL, what existing tools offer them, and the implications of these findings on future research and design.

Performance		
What Players Do/Want Help With During this Phase	***What Existing Tools Offer***	***Implications and Opportunities for Future Research and Design***
Determining an Opponent's Next Move, How to Respond, Information Tracking, How to Execute High-Level Skills, Non-Disruptive Aid, Beginner-Focused Aid	Predictions of Load-out or Victory, Recommendations and Instructions for Executing Strategies, Updates and Reminders for Tracking Information, Overviews of Past Performance	Dynamic In-Game Assistance, Emphasized Instructions, Monitoring Support, Non-Disruptive Assistance, Scaffolded Assistance

Table 7.7: A summary of what players do and want help with in the Performance phase of SRL, what existing tools offer them, and the implications of these findings on future research and design.

Self-Reflection		
What Players Do/Want Help With During this Phase	***What Existing Tools Offer***	***Implications and Opportunities for Future Research and Design***
Identifying Mistakes, Building Causal Relationships, Determining how to Adapt their Gameplay, Tracking Performance over Time, Obtaining input From Others, Avoiding Toxicity	Detailed Statistical Reflections, Retrospections depicting Gameplay Trends Over Time, Objective Evaluations to ID mistakes, Recommendations of Different Strategies, Online Databases for Easy Sharing	Safe Co-Regulated Learning, Emphasized Causal Understanding, Cyclical SRL through Forethought

Table 7.8: A summary of what players do and want help with in the Self-Reflection phase of SRL, what existing tools offer them, and the implications of these findings on future research and design.

Based on a comparison of the user requirements detailed within the interview and survey results against the taxonomy of existing interventions discussed in the previous chapter, I identified a number of explicit opportunities to improve computational support for SRL through research and design. These are summarized in Tables 7.6 through 7.8 and detailed below.

7.3.1 Forethought Phase (Before Play):

Detailed Strategic Predictions: Existing tools, during the forethought phase, offer predictions of what load-outs an opponent may select, including characters, positions, equipment, or skills. However, interview participants indicated that they would like predictions that go further, and offer information about what strategies an opponent may use. For example, players wanted more than just “the opponent is using character B they will probably be in X position” they wanted to know “because the opposing team has picked characters A and B they are likely going to use strategy X, which is characterized by maneuvers 1, 2, and 3, and time-points D, E, and F.” This detailed strategic information would help players derive more detailed plans themselves, as knowing more about what the opponent may do will allow them to better determine what they will do in preparation or response. For example, if they know the opponents are going to use strategy X, then they can plan to use a counter strategy. However, existing tools are far from offering this level of detail. While there exists research on strategic prediction [184], accurate, detailed predictions are still an open problem, making this an opportunity for future work.

Goal-Oriented Support: Goal setting and goal orientation processes are not only a critical part of the forethought phase of CPM and SRL, but also a central component of learning in general. In the context of games, previous work demonstrated the importance of goal setting to success [88] and found that setting and failing to reach inappropriate goals could lead to discontinuation of play. Further, the work I discussed in the previous part of this dissertation demonstrated that goal setting is a differentiating factor between expert and novice players [345]. Despite this, there was little emphasis on goal setting or orientation among the reviewed tools discussed in the previous chapter, and the interview participants did not explicitly discuss it either. E-learning tools designed to support SRL often emphasize goal setting during the forethought phase and have demonstrated that such an approach is beneficial to students [379, 299, 113]. As such, I recognize this as an opportunity for future work to explore, wherein it would likely be beneficial to players to include more explicit support for setting and monitoring progress toward gameplay goals.

De-Emphasized Review: We saw from the survey data and interview results that, while players saw benefits to reviewing past gameplay during the forethought phase, they were concerned about the potential negative effects of doing so. Specifically, participants suggested that being overtly reminded of what they had done wrong could leave them discouraged or in an otherwise negative state of mind, which would have an undesirable impact on their gameplay in the new game. Based on this finding, I suggest that future tool development consider ways to de-emphasize the review of past play during the forethought stage. However, I additionally recommend not removing it completely but investigating how to offer it in a way to help players without negatively affecting them. This may, perhaps, require more connection between the self-reflection phase and the forethought phase in existing tools, such that evaluations of gameplay can be brought into forethought but not necessarily happen there. This concept also better matches the processes of the cyclical phase model of SRL [756], which does not include reflective processes during the forethought phase.

7.3.2 Performance Phase (During Play)

Dynamic In-Game Assistance: Similar to the pre-game desire for more detailed predictions, players indicated a desire for predictions and recommendations that, during performance, updated to reflect the state of the game and give them relevant information on how to handle the current situation. Existing tools, however, rarely offered this level of support. Instead, existing predictions and recommendations were primarily static, based on known community trends, and not the state of the current game. As a result, players may receive a recommendation that does not actually make sense given the current state of play. For example, based on community trends, it might make sense to build a certain offensive item, and the tool may, thus, suggest the player do so. But if the enemy has a particularly powerful offensive character, it might make more sense to build a defensive item instead. Some research began to pursue dynamic in-game support [123], however, the topic is still under-explored. Thus, I recognize this as an opportunity for future work.

Emphasized Instructions: Players expressed an interest in instructions during the performance phase as a way to, especially for novice players, learn advanced skills. However, while instructions were present in existing tools, they were offered by only four tools, and each only offered them at one point in gameplay. This makes them the least utilized intervention out of the entire set discussed in the previous chapter. This is in spite of the fact that instructions are, perhaps, one of the best ways to help players learn to perform more advanced skills or maneuvers beyond what is presented to them by the tutorial itself. For those who may not belong to an established community that could help, such assistance is of particular interest and could be of great benefit. I thus recognize an opportunity for future work to improve players' experiences learning esports through a greater emphasis on instructional support.

Monitoring Support: Interview results indicated that players were not interested in using interventions to self-monitor during the performance phase, instead focusing almost entirely on strategy execution. The work I discuss in the previous part of this dissertation suggested that in-game interfaces are sufficient in supporting players' monitoring needs [345], and it may be that the current tools do not sufficiently provide anything new. Alternatively, previous work found that learners often do not self-monitor unless prompted to do so [333], and the same phenomenon may be occurring here. As such, I recognize this as an opportunity for future work. As performance monitoring has been recognized as critical to SRL [294, 756], it is exceedingly important to ensure that players are not only able to monitor but prompted to do so in ways that will improve their gameplay. Future design research can explore interventions to support performance monitoring in a way that the games themselves do not and that can prompt players to adopt this SRL process more extensively.

Non-Disruptive Assistance: E-learning tools designed to support SRL in the performance phase often give learners access to information about where they stand next to classmates or in relation to their goals, to help them make informed decisions about what to do next [294]. Esports, and many other complex games, however, are fast-paced, high-stakes environments, and, as the interview results illustrated, it can be difficult to implement assistance in a non-disruptive manner. Most existing tools either overlay assistive windows on the gameplay screen or require the player to look at a different screen, taking their eyes off the game in the process. That being said, interview participants were still interested in performance phase interventions. With this in mind, I recognize the design challenge of creating non-disruptive, performance-phase interventions as a potential direction for future research. Something to keep in mind while pursuing this goal, however, is accessibility. Some designers I have spoken with have discussed the implementation of audio-based assistance in order to reduce the amount of visual clutter on the screen. Such an approach, however, would be inaccessible to deaf or hard-of-hearing players.

Scaffolded Assistance: Interview participants suggested that in-game support can be of greater assistance to less experienced players but echoed the concerns of earlier participants who feared that they would be detrimental to skill development or more advanced play, effectively becoming a crutch that prevents the player from being able to play unaided. Previous work in learning has found that instructional techniques that are valuable to less experienced learners could very well be detrimental to the more experienced [325], lending credibility to these fears. With this in mind, I suggest that future design for esports assistants take inspiration from the scaffolding approach in learning [446, 67]. Under such a design philosophy, the assistant could dial back the level of support it offers as players master gameplay skills and gain expertise. Customizable configurations, in which players can manually indicate what they still do or do not wish to see, and when or how they wish to see it, would also make in-game support more palatable to advanced players. In doing so, future systems would evade providing unwanted input to skilled players, or building negative dependencies, while still being able to aid beginners.

7.3.3 Self-Reflection Phase (After Play)

Safe Co-Regulated Learning: I noted that the reviewed tools facilitated gameplay data sharing and its use for co-regulated learning [273]. However, survey respondents and interview participants indicated that, while sharing was beneficial, there were several dangers involved, namely the toxicity within the player bases [675, 676] and concerns that special strategies or unique maneuvers could be stolen. However, the advantages of learning from, and with the aid of, others are also well documented in the research literature [625, 273] suggesting that this type of support should not just be removed. Thus, I suggest that future systems explore ways to support co-regulated learning and cooperative learning through sharing of one's gameplay data while protecting the learner from negative interactions with other players. This may be realized through either anonymous sharing, sharing of only selected information, or through systems that allow players to share their data with only trusted others. While existing work on open community modeling is exploring this question [746, 599], this is still a new space with room for further investigation, especially as concerns over the ethical use of data become more apparent.

Emphasized Causal Understanding: The results indicated the importance of causality during self-reflection. Specifically, players indicated that it was crucial to identify how their in-game choices impacted their gameplay outcomes. While it seemed that players could glean this information from reflections, many emphasized that it was difficult to develop causal relationships given the design of existing reflection support systems. In other words, players often indicated that they knew when they did something wrong but rarely knew exactly what they did wrong. Further, even if they could identify their mistake, they struggled to know how to prevent it next time, out of an inability to know exactly what caused it this time. Based on this, I recognize an opportunity for future work to explore better ways to present causal relationships. One way to do this may be through the implementation and emphasized use of process-visualizations, such as those used in process mining, which present human processes in terms of state-action diagrams [686, 684]. I will expand on this idea further in the remaining parts of this dissertation.

Cyclical SRL through Forethought: The key to CPM is that the phases are executed repeatedly in a cyclical manner [750, 756]. This means that self-reflection and forethought should be connected to each other such that the results of evaluations and the chosen adaptive strategies that arise during self-reflection should directly inform new goals and plans during forethought. While a few tools offered recommendations during the self-reflection phase, which had the potential to link evaluations to future goals and plans, it seemed that there was a greater emphasis on the review of past gameplay during forethought, with many tools incorporating evaluations during the pre-game phase. As discussed above, many players felt that this could discourage them or cause them to dwell on their mistakes, effectively distracting them from performing better this time. Thus, I suggest that future work feature predictions or recommendations for future play during the self-reflection phase, as I believe it better emulates the processes inherent to SRL and better aligns with players' interests.

7.4 Summary:

In the context of this dissertation work, there are two primary takeaways. First, results indicated that review and evaluation-oriented processes engaged during the self-reflection phase were among the most common occurrences of CPM SRL among esports players and were among the most supported among esports tools. Second, results indicated that players are particularly interested in causality and being able to build relationships between their actions and the experienced outcomes. Further, results indicated that existing post-play visualizations are not sufficient for this. Based on this understanding, I arrived at the third thrust in my dissertation research, which looked into the potential of process-visualizations as a way to improve players' understanding of causality and elevate their self-reflection processes.

Part IV

Making Sense of Visualizations of Process

A prominent theme that emerged from the work described in the previous two chapters was the significant role that post-game visualizations of gameplay data played in helping players learn through self-reflection. The significance of reflection as a part of the gameplay experience is well documented [693] and many existing computational tools for complex games, especially those that emphasize data-visualization, acknowledge it as a critical element of learning [11, 701, 367].

However, the work in the previous two thrusts also illustrated how important it is to understand process and causality when trying to learn and improve at complex gameplay. Most existing tools [601, 459, 68] present data through aggregate visualizations. While effective at collating large amounts of data and making trends apparent at-a-glance, such visualizations are rarely able to preserve context or action-by-action information, which are necessary for understanding process and building causal relationships.

In this third thrust of my thesis, I explore visualization styles that present gameplay in a granular, action-by-action manner, thus preserving the gameplay process. These visualizations of process, which are primarily spatio-temporal [701, 11, 367], are prominent in the research literature. However, there is limited understanding of how players make sense of the, often dense, data they present. Thus, in this thrust, I explore meaning-making in the context of visualizations of process, looking specifically at two styles: spatio-temporal, and timeline-like process visualizations.

Chapter 8

An Interaction Taxonomy for Spatio-Temporal Gameplay Data

The work presented in this chapter was originally published in CHIPlay 2021 [1] [347]

8.1 Information Visualization and Interaction Taxonomies

Spatio-temporal visualization has grown notably prominent in games, as it allows human analysts to understand player processes by extracting them from telemetry data presented in low-level, granular, and context-sensitive ways [14, 705, 704, 232, 467, 706, 344]. User evaluations of spatio-temporal game data visualization systems, built both for players and developers, have demonstrated positive reactions from users, who find them aesthetically pleasing, easy enough to use, and sources of valuable information [683, 703, 274, 701, 706, 11, 367]. However, most existing evaluations focus on judging usability in order to identify design best practices. There exist few discussions of how users interact with and extract meaning from a visualization or use it to complete

[1]Kleinman, E., Preetham, N., Teng, Z., Bryant, A., & Seif El-Nasr, M. (2021). " What Happened Here!?" A Taxonomy for User Interaction with Spatio-Temporal Game Data Visualization. Proceedings of the ACM on Human-Computer Interaction, 5(CHI PLAY), 1-27. This research was led by me but would not have been possible without the assistance of Nikitha, Zhaoqing, and Andy and the guidance and insight of my advisor, Magy.

cognitive tasks. Further, those examples that do exist discuss these phenomena at a higher, abstract level, and often include this discussion as a supplement to other, higher priority results [366, 707]. As a result, detailed discussions of how users use, interact with, and make meaning from game data visualizations, especially in the context of learning, exist almost solely as theoretical frameworks [280, 80].

In contrast, the field of information visualization (infovis) has dedicated a great deal of time and effort to understanding the intricacies of how users perceive, make meaning from, and understand data. This knowledge of human understanding is formalized in what the field refers to as "taxonomies", or collections of behaviors and processes that users engage that should be considered when designing a new visualization. To date, many taxonomies and frameworks have been proposed, with new ones emerging each year [576, 736, 213, 576]. One of the most foundational examples, however, is Shneiderman's [613] "Task by Data Type" taxonomy, which identified seven behavioral tasks that users engage in when exploring visualized data. These were: overview (gain an overview of the entire collection), zoom (zoom in on items of interest), filter (filter out uninteresting items), details on demand (select an item or group and get details when needed), relate (view relationships among items), history (keep a history of actions to support undo, replay, and progressive refinement), and extract (allow extraction of sub-collections and the query parameters) [613]. Many visualization systems and more recent taxonomies explicitly or implicitly referenced Shneiderman's taxonomy in their development.

Although numerous taxonomies exist in the field of infovis to formalize interaction with data, the research area suffers from two notable weaknesses. First, very little work looks at the cognitive processes involved in making meaning through interaction with data, with most taxonomies focusing instead on the physical actions that users take to make sense of the data by making it readable or user friendly [598, 553]. This is a notable gap in the literature, as these cognitive processes are key to understanding how users turn what they can see in the visualization into actionable knowledge. Examples of the few taxonomies that do include cognitive elements include Yi et al.

[737], Yalcin et al. [731], Patterson et al. [506], and Valiati et al. [679] who focused their taxonomies specifically on these cognitive processes and included activities such as planning interaction and analysis [731], inferring [679], and managing mental models [737]. That being said, while Valiati et al.'s work validated the framework with a user study [679], these taxonomies are based primarily on literature reviews of existing taxonomies, rather than observation of users. This brings into question the extent to which they represent cognition in a comprehensive manner.

By taking a more human-centric approach, several visual analytics frameworks have been able to more prominently account for the cognitive side of sense-making and data interaction [191, 519]. Essentially, this domain proposes a "human is the loop" approach to visual analytics that aims to go beyond the interactive visualization of information, to better understand how human analysts make sense of data. A concrete example of work within this particular domain is the Knowledge Generation Model for Visual Analytics presented by Sacha et al. [572]. While not presented explicitly as an interaction taxonomy, this model includes cognitive elements, such as the formation of hypotheses and confirmation of those hypotheses through insights [572].

The second weakness of existing interaction taxonomy work is, similar to game data work, a notable lack of user studies, with the field instead favoring systematic reviews of existing visualizations and taxonomies. Notable examples of interaction taxonomies that are based on user studies include the work of Ziemkiewicz et al. [747], who studied how immunologists used data visualization in their lab work, and identified two main strategies (within graphs and between graphs), and Lee et al. [384], who derived the model of novice's information visualization sense-making (NOVIS) from a user study examining how people made sense of unfamiliar visualizations. However, the application areas for these studies are specific to the domain and population, and are therefore not easily generalizable.

As a result of the shortcomings described above, there is little consensus within the InfoVis literature regarding what elements a taxonomy should include. As a result, it is difficult to identify a single taxonomy that is appropriate to adapt to the domain of spatio-temporal game data. Further, most existing work focuses on visualizations of aggregate data, which may be applicable to aggregate game data visualizations, but not necessarily transferable to the granularity of spatio-temporal visualizations. As Rodrigues and Figueiras [563] demonstrated, applying the existing taxonomies to spatio-temporal visualizations can present a challenge, as it requires constructs to be reconfigured to account for the more granular data. Further, game data is exceedingly complex and existing taxonomies may not be able to account for this complexity. This suggests that the best approach would be for the domain of spatio-temporal game data to develop its own taxonomy, which is exactly what I set out to do in the first study within this thrust.

8.2 Developing a Taxonomy for Spatio-Temporal Gameplay Data

This first study sought to generate an interaction taxonomy for spatio-temporal gameplay data. Specifically, this study sought to answer the question "What are the interactive activities and cognitive processes that players engage in when they are analyzing spatio-temporal gameplay data?"

8.2.1 Methods

8.2.1.1 Defense of the Ancients 2

Similar to *League of Legends* (and arguably the game that inspired LoL) game of DotA 2 consists of two teams, the "Radiant" and the "Dire", both consisting of 5 players, who compete against each other in the same capture-the-flag style competition as LoL. Instead of a "nexus" the players seek to destroy the enemy "Ancient". The game map is similarly organized into three lanes, which house towers that fire at enemy players, with jungle areas between them. Both the lanes and the jungles are populated by various non-player characters that can be slain for experience and gold. Unlike LoL, however, players in a game of DotA 2 can also kill their own creeps, a move called "denial" as it prevents the enemy from getting last-hit gold bonuses.

8.2.1.2 Recruitment

Seven DotA 2 players (whose experience ranged from 3 to 8 years) were recruited via email, social media, and word of mouth from collegiate esports clubs and through snowball sampling. Participants were required to be (1) 18 years of age or older, (2) able to communicate in written and spoken English, (3) able to join a Google Meet call, and (4) in possession of enough experience with DotA 2 to identify and describe basic strategic behaviors.

8.2.1.3 Visualization Setup

Data from two DotA 2 games from a professional league were uploaded to the spatio-temporal visualization system *Stratmapper* [14], seen in Figure 8.1 (Please see the original paper for details of *Stratmapper's* functionality [14, 347]. I chose to use the data of others for three reasons. First, I wanted to observe cognitive processes of extracting meaning from the data, and felt this would be more apparent if the user did not already know the gameplay context behind the data. Second, I wanted all participants to observe the same set of events, which would not be possible with their own gameplay data. Third, interaction with game data often includes interaction with the data of others' with the goal of learning from their play.

Figure 8.1: The *Stratmapper* interface. At the bottom of the interface is the timeline, where the darker grey highlight is used to determined what data points are seen on the map (only the points encompassed by the highlight on the timeline will appear on the map). On the left is a list of heroes involved in the respective event. Clicking a hero's name in this list will mute their data points. On the left side of the timeline are buttons to mute data points for certain gameplay actions. Above the hero names is a drop down menu that can be used to change the visualized data set.

8.2.1.4 Think-Aloud Protocol

I created two types of tasks for this study. In one type, the participant was presented with a small, localized event and asked to recognize what was happening. I refer to this as a *recognition* task. In the other type, the participant was presented with a larger event and asked to identify three to five localized moments. I refer to this as an *identification* task. Each think-aloud session consisted of five tasks as follows:

- TASK 1 - IDENTIFICATION TASK: The participant was presented with a longer event that involved five heroes (three from one team and two from the other). The participant was prompted to find three to five gameplay maneuvers of their choosing in the visualization and describe them in terms of the actions, behaviors, and goals involved.

- TASK 2 - RECOGNITION TASK: The participant was presented with a short, localized event that involved three heroes (two from one team and one from the other). The participant was prompted to describe what event the data was depicting, again, in terms of the actions, behaviors, and goals involved.

- TASK 3 - RECOGNITION TASK: The participant was presented with a different short, localized event that again involved three heroes (two from one team and one from the other). The participant was again prompted to describe what event the data was depicting, again, in terms of the actions, behaviors, and goals involved.

- TASK 4 - RECOGNITION TASK: The participant was presented with a third short, localized event that again involved three heroes (two from one team and one from the other). The participant was again prompted to describe what event the data was depicting, again, in terms of the actions, behaviors, and goals involved.

- TASK 5 - IDENTIFICATION TASK: The participant was presented with a second longer event, this time involving all ten heroes on both teams. The participant was again prompted to find three to five gameplay maneuvers of their choosing in the visualization and describe them in terms of the actions, behaviors, and goals involved. This time, however, they were also asked to use *Stratmapper's* labeling system to apply labels to three to five gameplay events in the data. The purpose of this was to observe participants as they identified events , but with a persistent indication of what they had found so far (in the form of the labels).

Upon giving informed consent, every participant was shown the same *Stratmapper* instructional video and was given a chance to ask questions. Before beginning the think-aloud tasks described above, each participant was given an opportunity to practice using the tool and ask any clarifying questions regarding the visualization or the interface. Participants were then given instructions on how to think aloud and then prompted to begin the tasks. Each session lasted from 30 minutes to one hour and screen and audio were recorded.

8.2.1.5 Data Analysis

I worked with a collaborator and used ELAN [482], a tool for labeling videos, to analyze and code the recordings of the think-aloud sessions. We followed a thematic analysis protocol [237, 574], specifically focusing on how participants were interacting with the data, as observed from the videos, and the types of cognitive processes that the participants engaged in, as inferred from what they stated aloud. First we, separately, reviewed a subset of the data to develop initial lists of behavioral labels. We then reconvened and synthesized a combined list of labels through discussion and comparison. We then separated again and individually applied labels from the combined list to 30% of the data set [99], for an inter-rater reliability check. The resulting score, calculated using Cohen's kappa [133] was .82, indicating very strong agreement [375]. We then labeled the remaining dataset using the final list.

8.2.2 Results

8.2.2.1 Taxonomy

The taxonomy is organized into three categories: *Data Interaction*, *Sense Making*, and *Validation*. An overview of the taxonomy can be seen in Table 8.1.

Data Interaction	
Study Positioning to Construct Context	User studies the positioning of the players to construct an understanding of the setting for the data.
Study Movement to Infer Decisions	User examines movement over time and infers decisions based on their movement.
Seek Details to Support Sense-Making	User examines low level actions to obtain more context to explain data-points and make sense of observations.
Sense Making	
Leverage Domain Knowledge to Fill Gaps	User uses knowledge of the game to fill in gaps in the data.
Pinpoint Events to Frame Understanding	User identifies events of interest in order to build a frame for their understanding of gameplay.
Form a Hypothesis based on Context and Behavior that Evolves over Time	User provides or updates an explanation for what they believe is happening during gameplay.
Validation	
Review Events to Confirm Hypotheses	User ensures that pinpointed events conform with the chosen hypothesis.

Table 8.1: The taxonomy of user interaction for spatio-temporal game data visualization, consisting of seven activities organized across three categories.

Data Interaction: The three activities in this category relate to how users would use the tool to read the data. In other words, this category encapsulates observable interactions meant to facilitate the extraction of information and synthesis of meaning.

Study Positioning to Construct Context refers to the process in which a user constructs a mental model of gameplay context by examining and noting the locations of all of the players on the map. This activity specifically entailed the examination of stationary data points, representing gameplay actions, in terms of where they were located on the map, which indicated where a player was when they took the action represented by the point. These details, which indicated where players were and what they were doing, could be used to infer information about the context of play, i.e. which team was performing better or what point in gameplay the data was drawn from.

Figure 8.2: While studying positioning, participant 7 zoomed in on the middle lane area of the map, where the players of interest were located.

Participants would often make statements illustrating this. For example: “my guess is that both the Silencer and the Juggernaut here were pushing the creep wave here to attempt to break the tier three towers” (Participant 2) This demonstrates a moment during analysis in which the participant noticed that two Radiant side heroes were deep into the Dire side of the map and was able to make a judgment regarding the context of the game based on this positioning. Participants would also comment on the positioning of players in relation to actions, sometimes pointing out how these were odd or defied their understanding of context. For example: “it doesn’t make sense why Batrider is taking so much damage since he’s closer to his side of the map” (Participant 4). This demonstrates a moment during analysis in which the participant was confused about the greater gameplay context because they thought the presence of certain actions in relation to a player’s apparent position did not make sense. While studying positioning, participants would sometimes zoom in on an area of interest on the map for a closer and more granular analysis of the players’ positions. An example of this can be seen in Figure 8.2.

Study Movement to Infer Decisions refers to the process in which a user would infer players’ decisions based on their movement over time. Specifically, they would use *Stratmapper’s* interactive timeline feature to “scrub” backward and forwards in time, which would cause the data points on the map to shift position based on what actions were taken at the selected time and where, almost as if the data were animated. This, effectively, allowed the user to see the players' movement, from which they could make inferences about their decisions. Movement was sometimes studied in a holistic manner, with movement across the entire moment examined in one scrub-through from beginning to end. Sometimes it was also analyzed in a group manner, with multiple players visible simultaneously. Participant 7 is an illustrative example of this, who zoomed out enough to see the entire map and scrubbed through the entire moment, verbalizing their intent as: "I'm just going to look at an overview just to get an idea of what's happening." At other times, movement was examined more granularly, focusing either on smaller events within a moment or on a single player with other players being ignored or filtered out. This seemed to allow users to better understand the moment-to-moment decisions that were made by the player of interest. Examples of what participants would say regarding player decisions when studying movement are "...he rotates to the top lane through the river, turns back around for a second, but he keeps going" (Participant 1) and "you can see over here the Templar is chasing and hoping to dive the tower" (Participant 4).

Seek Details to Support Sense-Making refers to the process in which a user would support and facilitate the process of making sense of the data by intentionally seeking out more information regarding the state of the game. Specifically, seeking details is the act of seeking out low-level information that was not immediately visible from the tool, but which provided information regarding the state of the game. In the case of *Stratmapper*, detail-seeking manifested through interaction with the system's tooltips. Participants would trigger tool tips by hovering over icons in order to read detailed information regarding health, damage dealt, or who the target of an attack was. An example of a participant interacting with the tooltips to seek details can be seen in Figure 8.3.

I observed two patterns regarding how detail-seeking was used to help build a mental model of play. In the first pattern, *details were sought to answer a question* and detail seeking came after the other activities in this category, discussed above.

Figure 8.3: An example of participant 6 using a tool-tip to seek details.

For example, after engaging in activities in the data interaction category, participant 6 noticed that one hero was healing and brought up the tooltip to examine the hero's held items. When they found the answer they were looking for, they stated it out loud, saying "tranquil boots [a held item that heals over time], ok". In the second pattern, *details were sought to support the construction of context* and detail seeking often came before other interaction activities. For example, before engaging in data interaction activities in earnest, participant 5 wanted to know when during the game the data points occurred and brought up the tooltip to see the time stamp. They stated this goal out loud, saying "let's see, just to get a memory on the time stamp here...ok so a little bit later in the game". In both cases, seeking details supported the ongoing process of making sense of the data.

Sense-Making: The three activities in this category relate to the cognitive activities that users would engage in while interacting with the data in order to extract meaning and develop an understanding of gameplay.

Leverage Domain Knowledge to Fill Gaps refers to the process by which participants would use their own, pre-existing knowledge of the game to fill in gaps in the information provided to them by the visualization. There were two kinds of scenarios identified. In the first scenario, there was data missing regarding some aspects of the players' behavior, however, participants were able to use domain knowledge to fill in the missing information. For example, *Stratmapper* does not show ability usage, however, participants were able to leverage domain knowledge to infer ability usage based on observable information, i.e. "it looks like he surges [ability] away because he's running really fast" (Participant 3).

In the second scenario, information was included (usually within a tooltip), however, participants did not need to look at it to understand what was happening. For example, participant 1 was able to recognize that a player was farming a jungle camp based on their movement and did not need to look at who the targets of the player's attacks were, which was included within a tooltip. Participant 1 vocalized this during their think-aloud, stating: "when Oracle is pulling this creep camp here, you would have to read into it and you would see that that is what's happening because he's delivering damage to neutral camps, but I did not do that, I just assumed he was pulling based on his movements". Similarly, participant 5 recognized that a teleport was going to happen based on the positioning of a player and without having to check the items for a teleport scroll. They stated: "I already know that this is a pretty common, even in the modern updates of DotA 2, this is a pretty common spot to want to get away from someone, there's not a lot of vision in this area."

Players also levered domain knowledge to explain why certain patterns or details existed in the data. For example, participant 6 noticed, while examining tooltips, that a hero's health was increasing. After looking at the items the hero held, they were able to explain this pattern by using their own domain knowledge to supplement what was present in the data, stating: "tranquil boots [an item held by the hero in question] gives heal over time". At a higher level, all participants noticed a pattern in which data points for one hero, Legion Commander, were predominantly located in the jungles. All participants commented that farming the jungle was a common strategy for the hero, effectively leveraging their domain knowledge of popular gameplay strategies to explain an observable pattern in the data.

Pinpoint Events to Frame Understanding refers to a process by which participants frame their understanding of gameplay based on one or more easily identifiable or distinguishable events. These events were often pinpointed with the aid of *Stratmapper's* action icons. In some cases, these events were identified from a single action. For example, participant 7 pinpointed an event from a single death action, stating "and then there's a death on the Shredder". In other cases, these events were inferred from a collection of actions, often facilitated by an activity from the information interaction category. For example, participant 4 examined a set of stationary action icons at a given map location (studied positioning) and pinpointed a camp pull, stating "Oracle here, being the good support that he is, is pulling a neutral camp". There were also moments where an event could be pinpointed based on player positioning alone, without any action icons present. For example, participant 5 saw Juggernaut's location, during a point in time when the hero was taking no actions and pinpointed a setup for a gank, stating "Juggernaut is setting up [a gank] in the trees right here". In some cases, a pinpointed event would be the first thing a participant would analyze within a moment. For example, Participant 3 began their analysis of one moment with the following pinpoint: "it looks like Timber is going to die at the end, and no towers go down, and it looks like everyone else is going to be scuffling with some heals thrown in".

Form a Hypothesis of Context and Behavior that Evolves over Time refers to the process by which participants would provide theories for what was happening within the game, in terms of gameplay context or player reasoning or goals, based on what they could observe from the data. Three kinds of hypotheses were observed: hypotheses for what had already happened before the moment began, i.e. "just from the start of it, it looks like they may have already taken the tier two [referring to a tower]" (Participant 5), hypotheses for what had not yet happened but was about to, i.e. "it looks like Batrider is positioning, potentially, to jump that Slark" (Participant 3), and hypotheses for what was currently happening, i.e. "Juggernaut and Silencer, they might be pushing over here" (Participant 6). Hypotheses were also not entirely static, and participants would also form a hypothesis and then sometimes update it later as they discovered more information.

As can be noted from the three examples, a hypothesis was defined as a guess regarding game context, player goals, or player reasoning. In other words, it was an inference on the part of the user regarding information they could not observe, typically because it either was not included or because it was heavily dependent on information within the minds of the players, which could only be inferred. The formation of a hypothesis was facilitated by interaction with the data as well as leveraged domain knowledge, and a hypothesis could then be validated or invalidated by a pinpointed event. For example, participant 5 hypothesized that the hero Juggernaut "just flies over and snags him, I don't think there was a blink dagger", which was concluded by studying movement and leveraging domain knowledge. Specifically, they leveraged the knowledge that Juggernaut is able to jump to a target without the aid of an item and that an item called a blink dagger, which facilitates such movement, exists. However, participant 5 later pinpoints an event that overturns this hypothesis: "yep, ok, so he did blink in." Specifically, participant 5 was seeking details to confirm their hypothesis, saw the blink dagger in Juggernaut's items list, and was then able to pinpoint an event (the use of the dagger) that overturned their previous hypothesis.

Validation: This category consists of a single activity, which is categorized separately as it encompasses the act of checking rather than interacting with data or making sense of it.

Review Events to Confirm Hypotheses refers to the process by which participants were observed to validate their understanding of events at the conclusion of the analysis. Participants were observed to play through the entire scenario again, typically either stating or narrating what they were observing and how this supported their hypotheses. If everything they observed during this process supported their hypotheses, they would synthesize them into a single hypothesis for the entire moment and declare their analysis complete. Examples of this can be seen in: "so they gank him from both sides and he seems to die" (Participant 1) and "just a classical jump in and kill the carry" (Participant 5). If they observed data that contradicted one or more of their hypotheses, they would cease the review and re-engage in the other activities until they had satisfactorily updated their hypothesis. They would then review again.

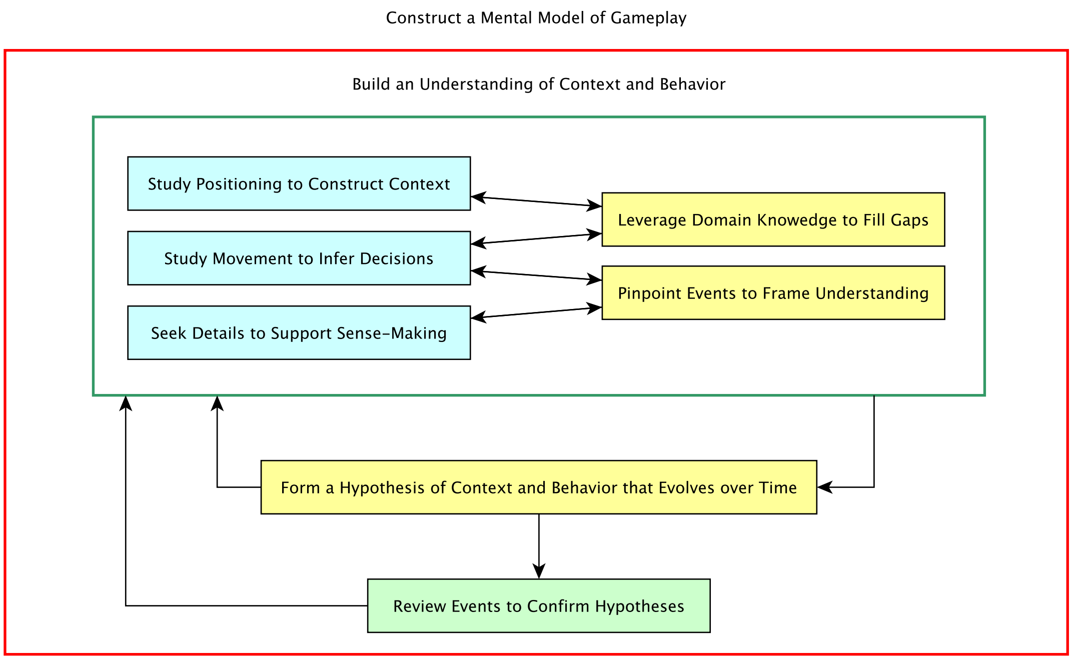


Figure 8.4: A preliminary process model to describe how users engage the activities in the taxonomy (see Table 8.1) in order to understand and extract meaning from spatio-temporal game data.

8.2.2.2 Process Model

By examining patterns in how the labels for the different action categories were applied to the recorded data in ELAN, discussed above, I identified a preliminary process model, seen in Figure 8.4, for how users make meaning from spatio-temporal game data. Specifically, users will engage in the data interaction activities of studying positioning to construct context, studying movement to infer decisions, and seeking details to support sense-making interchangeably with the sense-making activities of leveraging domain knowledge to fill gaps and pinpointing events to frame understanding. These activities are part of a larger process of building an understanding of gameplay context and player behavior from the data. That understanding is then used to form a hypothesis regarding context and behavior, which may evolve over time, and may trigger additional executions of the five previously mentioned activities. If the user believes they have enough accurate hypotheses to satisfactorily explain and understand the data, they will proceed to validation and review the events to confirm their hypotheses. If they find a discrepancy, they will then return to the process of building an understanding of context and behavior in order to form a more accurate hypothesis. This entire process is part of a larger process of constructing a mental model of gameplay, which is the ultimate result of interacting with the data, which allows players to understand why decisions were made and outcomes were experienced, and, ultimately, learn from the visualized gameplay data.

8.3 Implications for the Presentation of Gameplay Process

8.3.1 Positioning and Movement

Unsurprisingly, these results illustrated that users of a spatio-temporal visualization system spend the bulk of their time examining the positioning of data points and how that positioning changes over time (movement). However, it is the significance of that movement piece that is of interest here. Many existing spatio-temporal visualization systems do not include interactive features and therefore do not facilitate user control over the presentation of movement. Instead, movement is presented in a static manner, with arrows and paths drawn atop the map [701, 703].

The results of this study highlighted that movement and being able to observe movement, was critical to the overall process of interpreting spatio-temporal gameplay data and, further, demonstrated value in allowing users to control and interact with the presentation of that movement. All participants, at least once, studied movement by manipulating the timeline selection feature to see how the positioning of each player and their actions changed over time. This suggests that future spatio-temporal systems should consider including interactive features that allow users to examine movement in an animated manner. Further, users did not always go over the entire time frame of the moment, instead focusing on smaller events, going forwards and backward over them in an iterative manner. This result suggests that providing users with finite control, rather than a play/pause feature would better serve their needs.

8.3.2 Details on Demand vs. Domain Knowledge

One of the more prevalent elements of classic information visualization taxonomies, *details on demand* refers to the inclusion of granular, detailed information such that it can be accessed when needed [613, 576, 736, 731, 213] but also ensuring that it is only present when needed so as not to trigger cognitive load. During this study, we saw a variant of this concept manifest in the "seek details" activity, which was engaged by all users at some point, typically to confirm their hypotheses. This suggests the continued relevance of the concept and implies that future systems should continue to adhere to the practice, ensuring that all details are available within the interface and able to be accessed as needed.

However, a notable result of this study was the significance of domain knowledge, and the extent to which possessing domain knowledge meant that users did not need to seek details because they already knew what they were looking at. This emphasizes the importance of the "on demand" part of "details on demand" and implies that future systems may be able to reduce the risk of cognitive load or mental strain by hiding granular details from the main view on the assumption that the user may already know this information or be able to infer it unaided. This does, however, depend on the audience for the tool. While players are likely to be satisfied using their domain knowledge to interpret events, developers analyzing data in order to make game adjustments would likely want more information upfront. Further, those newer to the game, and with less domain knowledge, may require more details upfront as well as they will possess less domain knowledge. This suggests that variable arrangements of detailed information should be considered in future system design, depending on the skills, knowledge, and needs of the audience.

This also has implications for the abstraction of data. While abstraction is a promising avenue for increasing the readability of game data visualization [320], it may interfere with users' ability to trigger domain knowledge by removing too much information from what is immediately visible. Specifically, *Stratmapper* features no abstraction of data, instead presenting each data point as what it is, where, and when it occurred. This granular information was recognizable to the participants and triggered their domain knowledge, allowing them to generate an understanding of gameplay events. Abstraction may obscure too much information, and could possibly interfere with this process. Thus, future systems should be cautious of the extent to which data is abstracted, and may want to consider interfaces that include as much contextual detail as possible to facilitate the use of domain knowledge.

8.3.3 Pinpointing Events and Forming Hypotheses

Although the participants in this study were experienced DotA 2 players, they were examining data from gameplay that they were not a part of. Thus, a great deal of their ability to extract information from the data hinged on their ability to form a comprehensive hypothesis, which in turn, hinged on their ability to pinpoint events. These activities were at the core of mental model construction, the overall process that all of the activities in the taxonomy build into.

It is interesting to note that in infovis literature, mental models are often discussed in terms of how users perceive data or understand the functions of the visualization [337, 736, 731, 432, 398]. Here, however, the emphasis is on users building a mental model of the gameplay context, based on hypotheses of what occurred during gameplay formulated from the events they can pinpoint in the data. The iterative activities of forming and updating hypotheses are reminiscent of activities included within the NOVIS model proposed by Lee et al. [384]. Specifically describing how people make sense of unfamiliar visualization, the model includes "constructing a frame" to make sense of a visualization and "questioning the frame" when they doubt or need to verify it. I saw a similar pattern in how users generated a hypothesis and then sought verification through details and pinpointed events, especially if they came to doubt their hypothesis. However, in this case, the hypotheses were related to understanding the gameplay events depicted in the visualization, rather than the visualization itself.

It may be that the context-sensitive nature of gameplay data, and the need to be able to understand the context to make meaning of that data, prompted this behavior. It should be noted, however, that the gameplay events used in this study were unfamiliar, just as the visualizations in Lee et al.'s study were unfamiliar [384]. Thus, it may be that the iterative acts of creating, validating, and updating hypotheses are driven more by a need to understand unfamiliar artifacts than any specific domain context, and future research may want to explore this further.

For future systems, the key roles pinpointing events and forming hypotheses play in building a mental model of gameplay suggests opportunities for models or AI assistants that can aid users by automating the processes, perhaps by automatically identifying key events and drawing the user's attention to them or by proposing a hypothesis for the user to edit if necessary. Given that user's hypotheses were often updated after being first proposed, the results also emphasize that accurate identification will not be an easy task, and that human users should always be able to change, update, or correct any machine-driven identification. Further, the emphasis of these activities also implies the value of tools such as *Stratmapper's* labeling system, which can allow users to track what they have identified as they continue their analysis, and allow them to better remember their hypotheses when they review.

8.3.4 Getting the Big Picture vs. Going Granular

From the results, both the taxonomy and the process model, it is apparent that interacting with spatio-temporal gameplay data involves both holistic and granular perspectives of the data. Specifically, as discussed in the results, participants would often begin by examining the positioning and movement of everyone across the map, to get an overview of events. Then, they would get into the details of specific maneuvers, zooming in and focusing on specific players. This resonates with what is discussed in existing taxonomies, which discuss gaining an overview first [737, 576, 613], then exploring data more granularly by zooming and filtering [213, 731]. For future systems, this suggests that they should ensure that users have the ability to see data in both a holistic manner, in which they can see everything at once, and a focused manner, in which they can zoom into an area or player of interest.

When discussing these insights though, it is noteworthy that participants did not always engage in filtering, despite this being an action described extensively in previous work [613, 576, 736] and a feature supported by the *Stratmapper* interface [14]. It was because of its inconsistent, and not predominant, use that it was not included as a unique activity within the taxonomy. Those participants who did use *Stratmapper's* filtering mechanism would only filter players out temporarily, usually to facilitate a more granular study of a given player's movement patterns, and would always bring their data back to the visualization afterward. Further, no participant ever filtered out any type of event, despite this being a functionality of the tool. This aversion to filtering, and tendency to restore the unfiltered data, suggests that participants were reliant on the additional data to make sense of what they were seeing. It may be that the additional data points played a key role in building contextual inferences and hypotheses and that users were, therefore, hesitant to remove them our of concern that they would miss events critical to comprehending gameplay. This may be particularly relevant in the context of team based, multiplayer games, as it is often the actions of one player that influence the actions of another. Future work may want to examine this phenomenon in further detail, to examine the extent to which filtered vs. holistic data is able to impact users' understanding of events.

The act of undoing filtering and restoring the previous view within the visualization brings to mind the concept of "history", or the ability to return the visualization to a previous state, discussed in existing taxonomies [613, 213, 576]. However, while previous work discusses this in the context of retracing one's steps or undoing a mistake, here it appears to be more of a case of restoring the big picture when one has completed granular analysis. For future systems, this suggests that, although the ability to go granular with analysis must be supported, it must not be done in a way that removes data from the user's view permanently. Any filtering that is facilitated must be able to be removed, such that the context of the game at large, as inferred from the extra data, can be restored. This need to restore the big picture is further supported by the "review" activity, which was always enacted with no player filtered out. Thus, it would appear that it requires the presence of all of the data such that the user is able to check their theories and ground the findings of their granular analysis within the context of the gameplay as a whole.

8.4 Summary

Based on these results, I obtained valuable insight into the ways players will leverage process information towards making sense of gameplay. Specifically, there is a strong emphasis on understanding decision-making and reasoning. However, this work looked at spatio-temporal visualizations, which are not applicable to every type of complex game, as they are dependent on the inclusion of a spatial environment that can be navigated. Thus, in the context of this dissertation, I can build on this work by turning our attention to what I call process visualizations, which do not require the game to have a spatial component and are therefore a viable option for presenting players with granular process information across multiple genres and game types.

Chapter 9

An Interaction Model for Process Visualizations

The work presented in this chapter was originally published in ICEC 2022 [1] [350].

9.1 Visualizations of Process in Games and Beyond

Recognizing that complex games need a process-oriented way to present data to players, such that they may learn from it, that does not require the game to have a spatial component, I turn to other domains that have emphasized the depiction of process for inspiration. Specifically, I turn to process mining. The domain of process mining has, for years, been focused on capturing and understanding human processes, or sequences of decisions and actions, such that systems and workspaces can be better designed to support them [686, 685]. Within the domain, processes are mined from event logs of interaction with a system in order to better understand humans and better

[1]Kleinman, E., Villareale, J., Shergadwala, M., Teng, Z., Bryant, A., Zhu, J., & El-Nasr, M. S. (2022, October). Towards an Understanding of How Players Make Meaning from Post-Play Process Visualizations. In Entertainment Computing–ICEC 2022: 21st IFIP TC 14 International Conference, ICEC 2022, Bremen, Germany, November 1–3, 2022, Proceedings (pp. 47-58). Cham: Springer International Publishing. This research was part of an NSF-funded project conducted collaboratively with the PXL lab at Drexel University and ITU Copenhagen. It would not have been possible without the assistance of Jen, Murtuza, Zhaoqing, and Andy and the guidance and input of Jichen and Magy.

design systems to facilitate and support them [685]. Understanding the mined processes, however, is dependent on being able to read the output in a detailed, comprehensive, and interpretable manner.

In order to do this, process mining, as a field, has developed a series of visualizations to help managers, analysts, and designers glean processes from event logs [551, 718, 687]. These visualizations all follow a similar approach, using node-link diagrams to display the ordering and progression of human actions as they work towards the completion of a task [514, 718, 687, 581]. The differences exist in how each visualization style depicts complexity in the process. For example, Petri Nets emphasize state information in addition to actions taken [514, 551], while Causal Nets depict the existence of two potential actions that can be taken at the same point in the process through “or” nodes [687].

In games, graph-based process-visualizations exist almost exclusively in the context of game analytics and user experience research [187, 344, 297]. Glyph [478] is an example of one such process visualization, which uses a node-link diagram to display player processes in terms of state-action transitions, where the state is the state of the game and the action is what the player did that moved them to the next state [478]. Glyph has been used to extract strategic trends from educational puzzle games [322, 478], esports [14], and even an augmented reality game [320]. Other tools for analysts and researchers have also leveraged the node-link diagram approach, including Playtracer, [26] and Play-graph [702].

In the context of player-facing tools, process-visualizations are typically designed as timelines depicting the ordering of actions taken over the course of a game [347, 367, 11]. For example, Kuan et al.’s system for Starcraft 2 includes a timeline, a process visualization, of the order that structures were built in [367]. Visualeague, a system for League of Legends, similarly uses a timeline to present the order in which the player leveled up their skills [11]. Even *Stratmapper*, discussed in the previous chapter, featured a process-visualization in the form of a timeline [14, 347].

However, these timelines are often secondary features attached to a spatio-temporal, map-based visualization system, as they are in all of the above examples [367, 11, 14]. In fact, more recent iterations of Visualeague have de-emphasized the process visualization and focused more on the spatio-temporal data and aggregated stats over time [10]. Being relegated to a supplementary visualization means that player-facing process-visualizations are not often the focus of research. As such, there has been little work exploring ways to expand these visualizations to present more comprehensive gameplay information or exist as stand-alone tools.

This is a gap worth exploring, given that the work I discussed previously emphasizes the significance that being able to understand process has on SRL, specifically on self-reflection, in the context of complex games. While spatio-temporal visualizations are able to address this, process-visualizations may offer a viable alternative that better generalizes to all types of games. Thus, in this next project, I sought to develop a stronger understanding of how meaning is made from data presented in a process visualization of gameplay, in order to better inform future exploration of the design.

9.2 Making Sense of Process Visualizations for Games

This study sought to expand on the previous one by looking at meaning-making in the context of the aforementioned process visualizations and answering the question "What interpretation techniques do players use to extract meaning from sequential-process visualizations of others' gameplay?"

9.2.1 Methods

For this study, I used the game *Parallel* [745], an educational game that teaches parallel programming. While not an esport, it still qualifies as a complex game due to having multiple correct solutions and unpredictable outcomes.

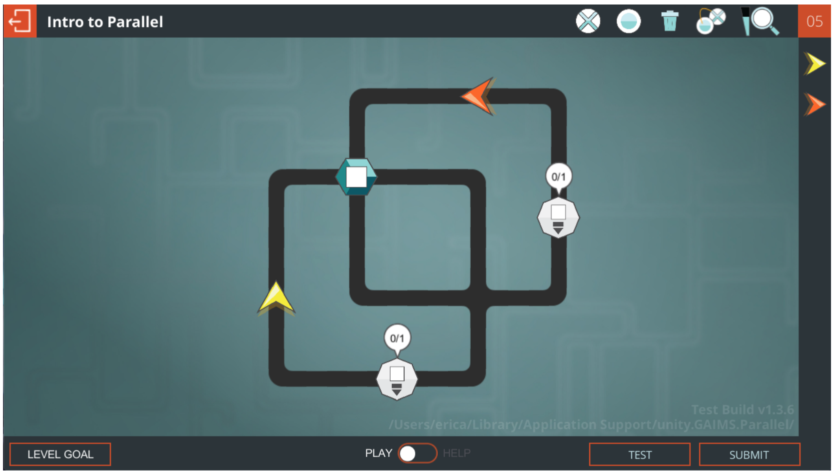


Figure 9.1: A screenshot of the game *Parallel*. Players need to place semaphores and signals to direct arrows, which carry packages and move along pre-defined tracks, to the designated delivery points. The player must coordinate threads executing in parallel. The level pictured was the subject of the retrospective interviews.

9.2.2 Parallel

Parallel is a single-player, 2D puzzle game designed to teach concurrent and parallel programming concepts [745]. The game introduces students to concepts such as non-determinism, synchronization, and efficiency using visual representations (see Figure 9.1).

The player's goal is to coordinate the arrows, which represent parallel programming threads and move at random speeds, to move packages from their pick-up points to their designated delivery points. To accomplish this, players must place signals and semaphores on the track and link them to control each thread's movement. Semaphores, represented by the circle with an X in the upper right corner in Figure 9.1, will block the movement of an arrow unless opened by a signal, represented by the other circle in Figure 9.1. When a player thinks they have a correct solution they can test or submit it. A test will check the solution with one possible speed of arrow movement, while submission will test them all.

Action	***Definition***
Place Semaphore	The player places a semaphore on the board
Place Signal	The player places a signal on the board
Link Signal and Semaphore	The player links a signal and a semaphore
Test Passed	The player runs a test and it passes
Test Failed	The player runs a test and it fails
Stop Test	The player stops a test simulation before it completes
Stop Submission	The player stops a submission simulation before it completes
Toggle Semaphore	The player locks or unlocks a semaphore
Move Semaphore	The player moves a semaphore to another spot
Move Signal	The player moves a signal to another spot
Destroyed Semaphore	The player destroys a semaphore
Destroyed Signal	The player destroys a signal
Submission Passed	The player submits a solution and it passes
Submission Failed	The player submits a solution and it fails
View Help	The player views the help menu

Table 9.1: This table showcases all 15 in-game actions used to analyze the players' gameplay.

9.2.2.1 Recruitment

13 undergraduate computer science students were recruited. Participants were required to be (1) 18 years of age or older, (2) located in the United States, (3) able to communicate in written and spoken English, and (4) able to play on a Windows machine. Participants did not have to have prior knowledge of parallel programming. Study participants played four levels of *Parallel* including the one shown in Figure 9.1.

9.2.2.2 Visualization Setup

To generate the process-visualizations, I worked with collaborators who helped design the game and identified a set of gameplay actions necessary for recognizing strategic processes (see Table 9.1). A Python script was prepared to filter and format the log files into sequences of these actions. Gameplay actions were abstracted from lower-level logged events by the script. For example, "place semaphore" (a gameplay action) was abstracted from a pair of logged events: "mouse click down" (over semaphore icon in the

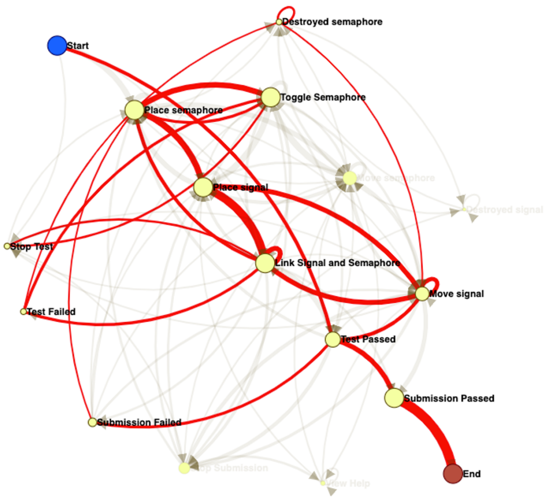


Figure 9.2: An example of the process visualization used in the Parallel study.

menu) and “mouse click up” (over a track). Any logged events that did not correspond to one of the selected gameplay actions were filtered out.

Using the visualization tool Glyph [478], I generated process-visualizations for each player using the 15 key actions, an example of which can be seen in Figure 9.2. I chose this tool as Glyph’s network graph resembles other prominent types of process-visualizations, such as dependency graphs [759], those used to visualize process mining results [686, 684], and customer journey maps [663, 566], making it a representative example of such a visualization. At the same time, previous work has demonstrated its generalizability, scalability, and usability with game data [14, 478, 322]. Please see the original paper for further details on *Glyph’s* functionality [478].

Glyph’s network graph uses a node-link diagram to represent player behavior as a process visualization. Each node represents a different in-game action and a link between two nodes indicates that at least one player transitioned between those two actions. As a player repeats actions during gameplay, their trajectory will loop back to already visited action nodes. In other words, if a player [placed a semaphore], [placed a signal], and then [placed a semaphore] again, their trajectory through the visualization would go from the [place semaphore] node to the [place signal] node and then back to the same [place semaphore] node. The thickness of a link indicates how many players made that transition (with thin being less and thick being more). Individual player trajectories within this visualization can be highlighted as seen in Figure 9.2. Each action node, combined with the varying thicknesses of links and the ability to highlight specific players, allows the visualization to scale to community data while maintaining both process information and readability [478, 14, 322].

9.2.2.3 Interview Protocol

Based on previous work, I recognized that interpretation techniques tend to fall into two categories: interaction techniques used to extract information from the visualization and cognitive techniques used to make sense out of that information. In order to ensure that I elicited techniques in both groups, I developed two prompts:

1. Could you describe what you understand about this player's actions from this visualization?
2. Can you say a bit about why you think the other player played the way they did?

I refer to the first prompt as the "interaction prompt", as the primary goal of this prompt was to elicit interaction with the data and the extraction of information. I refer to the second prompt as the "cognitive prompt", as the primary goal of this prompt was to encourage cognitive processes and sense-making by asking participants to think about the other player's reasoning behind their actions.

To focus the data collection on how participants read the data and ensure that interaction affordances, such as zooming, would not become confounding variables, a slide deck was prepared to present to each participant during the interview. One researcher led the interview, screen-sharing the slide deck, while two others remained silent and recorded, in text, what the participant said.

The first slide contained a visualization of the participant's own data. On this slide, the lead researcher gave the participant basic instructions on how to read the visualization. The next slide contained a visualization of another participant, who played similarly to the interviewee. The last slide contained a visualization of another participant, who played differently than the interviewee. While displaying the second and third slides, the lead researcher asked the prompts described above. Interviews lasted about 30 minutes and participants received a 50$ gift card.

9.2.2.4 Data Analysis

Interview data were analyzed using a two-step, iterative thematic analysis protocol [237, 574]. The first step of the analysis sought to identify the specific, individual interpretation techniques that players used when making sense of the data. To do this, myself and a collaborator, separately, analyzed 30% of the interview responses. This generated a combined code book containing six interpretation strategies. We then separated and performed an inter-rater reliability validation using Cohen's Kappa [133] on 30% of the data. The codes achieved an IRR score of .87, indicating very strong agreement [375].

The second step sought to identify the overall process of making sense of the data. Here, we separated again to analyze 30% of the data. Again, the unit of analysis was an entire response to a prompt. We analyzed each prompt response and marked what techniques were apparent in the response and in what order they were used by the participant. We then reconvened and discussed and identified two unique methods for engaging the interpretation techniques across the data set. We then separated again and performed a second inter-rater reliability validation for the two methods, again on 30% of the data. The method codes achieved an IRR score of .74, indicated strong agreement [375]. I then labeled the remainder of the data set with the method codes.

9.2.3 Results

9.2.3.1 Interpretation Techniques

The types of interpretation techniques and their definitions can be seen in Table 9.2.

Reading the Visualization to Collect Information refers to the user reading the presented data but not offering any insight or meaning behind it. It appeared as though participants would engage this technique when they were trying to collect information from the visualized data as a precursor to making connections between data points. For example, Participant 9 reads the trajectory of another player, stating "They do start, test passed, sub failed, they place the semaphore, then maybe they toggle it, they place the signal, they link, maybe they move it around."

Notably, reading the visualization would often encompass a read-through of the entire sequence, suggesting that participants were engaging this method to gain a holistic overview of the data. This is illustrated by Participant 0 who said "it looks like they placed and moved semaphores and placed signals, linked some together, ran a submission, stopped the submission, placed another semaphore, maybe moved it again, toggled it, and then maybe toggled a different one, ran it, a test passed, and then the submission passed." This is reminiscent of an interaction paradigm often discussed in information visualization taxonomy literature, where users will begin with the big picture, then zoom in on areas of interest or look for connections between data-points [576, 613]. This leads into the next technique we describe, in which participants, armed with the big picture, would begin looking for connections between data-points.

Participants did not always read the visualization, with many jumping right to identifying patterns. This is likely the result of individual differences, with some people adapting more quickly to the visualization style and being able to more quickly extract points of interest.

Identifying Patterns to Inform Inferences refers to the user would identifying and extracting a pattern from the data. Participants would make general statements about the characteristics of the data or extrapolate on the visible information in some way. It was through this particular technique that we saw participants begin to make connections between data points. Further, we saw participants use pattern identification as a way to begin forming inferences about the players who produced the data. We saw two types of patterns that participants would identify in our study:

- **Sequential Pattern**: Refers to the participant identifying patterns in the ordering of actions. For example, Participant 3 noticed that "[the other player] repeats that process of toggling then placing then linking" (Participant 3). Recognition of these patterns is facilitated by the sequential nature of the visualization and it would likely be harder to recognize such patterns without them. We also noted that being able to identify sequential patterns helped participants derive inferences about the thought processes of the players who produced the data. For example, Participant 8 described "They seem to jump back and forth a lot. They were probably thinking through a lot of their placement and movement." The first part of this statement, in which they recognize a pattern in ordering in which the other player often jumped between toggling and movement, goes on to inform an inference about the player's reasoning, that they were thinking through their solution.

- **Frequency Pattern**: Refers to the participant identifying a pattern regarding the number of actions taken. For example, Participant 0 said: "They move a limited number of times, but they ran the submission a lot because it looks like they stopped it a lot." The emergence of this technique may be informed by the nature of the visualization, as it indicated how many times an action was taken. Further, it is unlikely that this pattern is unique to process visualizations, as aggregate visualizations also display frequency information, and often more clearly. This is in contrast to the sequential pattern identification technique, which is likely better facilitated, as well as encouraged, by a process visualization. We also observed that recognition of frequency patterns would lead to inferences regarding the player who generated the data. For example, Participant 7 said: "It looks like they placed a lot so they probably deleted them instead of moving them". We can see from this example that recognition of the frequency pattern, that elements were placed a lot, resulted in an inference about the player who produced the data and how they approached gameplay.

Pattern recognition is a critical part of the process of making sense of visualized information, and is frequently discussed in previous work [576]. Of interest, however, is that the patterns informed and were accompanied by inferences about the player who produced the data. This is similar to the phenomenon observed by Kleinman et al. in their study on spatio-temporal visualization [347]. We discuss inferences, as an interpretation technique, in more detail below.

Making a Comparison to Guide Pattern Identification refers to the user comparing the highlighted player's data and their own experience. Comparison did not always occur, but when it did, it was typically connected to identified patterns. For example, Participant 0 said "They use the stop submission button, that's interesting, I don't think I used it at all." Often, participants would use comparison as a way to guide the identification of additional patterns. This is well illustrated by Participant 1 "Once they laid down a solution they would test it and see if it failed or not. Whereas I don't remember doing as much testing." Here, the participant has identified a pattern in which the subject would lay down a solution then test it. They compare this to their own gameplay, in which they did not test as much. This understanding that the other player took a different approach can act as a preliminary guide to help them identify more patterns (what else did the other player do differently?) and begin to generate a more formal inference.

Technique	***Definition***
Reading the Visualization to Collect Information	The participant read the data that appeared in the visualization but did not extrapolate on it
Identifying Patterns to inform Inferences: Sequential Pattern	The participant identified a pattern related to the ordering of data-points
Identifying Patterns to inform Inferences: Frequency Pattern	The participant identified a pattern related to the amount of data-points
Making a Comparison to Guide Pattern Identification	The participant compares the data of the other player to their own experience to better make sense and extract patterns
Making an Inference to Understand the Other Player: Approach or Strategy	The participant makes an inference regarding the subject's intentions behind the actions they took
Making an Inference to Understand the Other Player: Understanding or Expertise	The participant makes an inference regarding the subject's knowledge or comprehension of the gameplay

Table 9.2: The six interpretation techniques identified based on analysis of players' interaction with the community visualizations and brief definitions.

Making an Inference to Understand the Other Player refers to the user making an inference about the subject who produced the data. As discussed above, inferences were informed by identified patterns within the data, which were sometimes guided by comparison. While inferences are known to occur in infovis interaction [576], Kleinman et al. [347] illustrated how game data visualization varied in that the inferences are about the intentions of the players rather than the data itself. Here, however, we noticed a difference from Kleinman et al.'s work. They observed participants often making inferences about game contexts. In contrast, we observed inferences to be focused primarily on the player and their own decision making processes. This is likely informed by the visualization. Whereas Kleinman et al. used spatio-temporal visualization, therefore displaying more context and likely encouraging participants to focus on it, the process visualizations we used displayed only the player's actions, encouraging participants to focus solely on the player and their reasoning. We observed two types of inferences from our participants:

- **Approach or Strategy**: Refers to the participant making an inference about what the subject was planning and how they were executing that plan. For example, Participant 7 said: "It looks like they were checking their work a lot as they went. So this would help them see how the changes they made would impact the final result, visually". Through these inferences, subjects attempted to make sense of the data, and more specifically the patterns they had identified, by understanding the reasoning and intentions of the other player. Another example of this comes from Participant 10, who said "They saw that the test passed so their aim was to try and generalize the solution." Here, participant 10 infers a strategic decision that the player made (trying to generalize their solution) as a way of explaining an observed pattern in the data (that the player did not immediately submit after their test passed and instead took other actions). In doing so, the participant develops an understanding of the other player that allows them to make sense of the data.

- **Understanding or Expertise**: Refers to the participant making an inference about what the subject knows about the task or subject. For example, Participant 4 said "It looks like they weren't so sure how to lay everything out since they kept moving around". By suggesting a level of knowledge, these inferences gave participants an understanding of the data based on an image of the other player, specifically focused on what they understood. For example, Participant 11 said "I would say they probably came in with a good idea about how they were going to do the level before they started playing [since] they're very calculated, they rarely jump back and forth between states." Here, the participant has developed an image in their mind regarding the expertise of the other player (that they had a good idea of what they were going to do) that can be used to explain an observed pattern in their data (that they rarely go back to previously visited states).

It is likely that the emphasis on making inferences to understand the other player were encouraged by the emphasis on a single player's trajectory within the visualization. However, being able to single out and understand the decisions of a single player is a key element of learning. Open learner models [294] demonstrated that providing users with access to someone else's data, can motivate them to spend more time in the system and help them locate relevant material. Other work has supported using OCMs to help users to gauge their relative progress towards a shared goal [86], effectively seek collaborators [90], and even follow in the footprints of other, more successful, users [85].

9.2.3.2 Sense-Making Methods

The interpretation techniques connect to one another to form a process for making sense of the data. We refer to this process as a sense-making method. However, while the results discussed above illustrate how reading the visualization informed pattern identification and comparison, which in turn informed inference making, we observed, in our analysis, that there were times in which participants would begin their sense-making by making an inference, often informed by the surface-level details of the visualization (such as the length of the highlighted trajectory). Thus, we identify two general methods for sense-making for retrospective process-visualizations for games, described in detail below:

9.2.3.3 Induction Method

: This method represents an approach in which the players were observed to begin their sense-making process by reading the visualization. They would then identify patterns in the data, either sequential or frequency, and use comparison if necessary to generate an understanding of the gameplay events represented by the data and guide their analysis. This would culminate in an inference about the other player, either about their approach or their expertise. In more general terms, participants who used this method would begin by collecting information and generate an inference about the player that could explain what they saw in the data.

An example of this method is demonstrated by Participant 7: first, they read the visualization, stating "Ran a test and it passed then worked to place the items in one sequence, and then the test failed, and then in another they stopped it again." They follow this with recognition of a frequency pattern, stating "It looks like they placed a lot". Finally, they offer an inference of the player's strategy, meant to explain the information they have collected, stating "they probably deleted [the signals and semaphores] instead of moving them."

An example of comparison being used in this technique is provided by Participant 11. First, the participant noted a frequency pattern, stating "It seems like they move their semaphore a lot," which they followed with a comparison, stating "Which is something I didn't do." They then arrived at an inference of the other player's approach, stating "I think they didn't form their idea before placing things and needed to change it." Participants who used this method ended their analysis when they had collected enough information to make an inference, and thus the method had a clearly defined endpoint.

9.2.3.4 Framing Method

: In this method, participants begin with an inference, before proceeding to information collection. When participants used this method, they would first make one or more inferences about the other player's approach or understanding, often based on visually apparent details such as the length of the trajectory, which were then used as a framing device for making sense of the data. They would then switch to collecting information, first reading the visualization, then using one or both pattern identification techniques and comparison, if necessary, to generate hypotheses that justified and supported their initial inference.

An example of this method is demonstrated by Participant 12: they begin with an inference of the other player's strategy (or lack thereof), stating "I would think that this player kind of did stuff at random, I'm not sure if there was a process that they used.” They follow this by reading the visualization to collect information, stating “It seems like [they’re] going from start and then placing a semaphore [then] going from test passed to stopping submission and moving a signal”. They follow this with identification of a sequential pattern (or lack thereof), stating “It doesn’t look like this graph had a lot of iterative processes...It’s a little jumbled up.”

An example of comparison being used with the framing method is demonstrated by Participant 11: they begin with an inference of the other player’s understanding, based on their shorter trajectory, stating “They had a clearer sense in their mind of what they wanted to do.” This is followed with observations of patterns in the visualization, where they state “the nodes aren’t overlapping too much, the thickness stays about the same.” Their understanding of these patterns are framed by a comparison to their own experience, stating “it also seems like they took a different approach from me where they placed the semaphore then the signal and linked those whereas I put two at the same time.” Participants who used this method seemed to end their analysis when they felt they had collected enough information to sufficiently support their inference. While making an inference and then seeking out data to support it has been observed before in information visualization [384], it is noteworthy here that we observed no instances of a participant changing their initial inference.

9.3 Guidelines for the use of Process Visualizations in Complex Games

It is apparent from these results as well as those discussed in the previous chapter that inferences facilitate players’ ability to extract actionable insights from data. As I said before, this finding is similar to what has been discussed in InfoVis work regarding mental models of data [736, 737, 384]. Unlike InfoVis work, the inferences here inform a mental model of the individual who produced the data rather than the data itself, similar to what I saw in the previous study [347] but different in that there it was more focused on the context of the game where here it is more focused on the player themselves. Further, in this work we see the presence of a sense-making method that begins with the inference and then collects data to enforce it. This may have been encouraged by the nature of the visualization, from which surface-level information, such as length of trajectory, could be quickly extracted and used to reach a preemptive conclusion, whereas in the previous chapter, it was more difficult to extract such quick judgments from the spatio-temporal data.

This suggests that process-visualizations, which present data in a holistic manner, may encourage players to make assumptions about the data upfront. However, there is a very real possibility that these up-front assumptions can lead to inaccurate inferences. A possibility of particular concern given that, unlike the with the previous chapter, participants in this study were not seen to update or adjust their inferences. **Thus, process-visualizations for post-play analysis should consider incorporating design elements that can inform players' up-front assumptions and guide them towards correct initial inferences.** One way to accomplish this could be grouping or labeling actions inside a visualization to indicate what they mean.

As I stated above, participants in the study who used the framing method were not observed to make adjustments. It seemed that they rarely uncovered information that they recognized as contradictory to their inference, a sharp contrast to the participants of the previous study. Based on these results, I hypothesize two explanations. The first is related to the participant's familiarity with the game. In this study, participants had no prior experience with *Parallel*, as opposed to the previous study where they were experienced DotA 2 players. As a result, they may lack the domain knowledge necessary to recognize gameplay strategies in the data that would contradict their hypotheses. **Thus, process-visualizations may aid players best if they are not displayed until the player has become more familiar with the game.**

The second explanation is related to the abstraction of the data. The presentation of the gameplay data as a trajectory of actions may have been too abstract. Including game state information, which was present in the spatio-temporal visualization used in the last chapter's study, in the process visualization could have helped players better understand what they were observing. **Thus, retrospective process-visualizations should consider incorporating game state information, to ensure that players are able to correctly interpret the context behind each action.** This implication, along with the previous one, can help ensure that the player is equipped to correct misunderstandings about the data.

Additionally, the inclusion of comparison, as shown in the results, did not emerge from the previous study, where players were not shown their own data. This suggests that the inclusion of a player's own data is likely to spark comparison between themselves and others. Using comparison between self and others has been explored in the domain of personal informatics, though usually within the context of a user understanding their own data through the comparison [539, 192]. Here the comparison was used to understand the other player, as finding the differences in how the other player behaved compared to oneself gave participants an anchor point to begin understanding the rest of their experience.

This suggests that process-visualizations can leverage comparison to help players more quickly identify connected patterns and reach inferences in a space where excessive data is not available. **Thus, process-visualizations in post-play contexts should consider highlighting how the player's own data compares to and differs from the data of the subject of analysis.** This does, however, raise questions about the potential risks of prompting comparison among players, as previous work has demonstrated that players who under-perform can become discouraged when prompted to compare themselves to high-performing players [196]. Thus, process-visualizations may wish to only permit comparison against other players with similar skill levels or quality of performance.

9.4 Summary

This study built on the previous work by adjusting our understanding of how players make sense of process data and revealing guidelines for the presentation of said data via process-visualizations, which, like spatio-temporal visualizations, present granular, action by action, data but do not require the game to include a spatial component, thus proving more generally applicable. These highly generalizable visualizations, I argue, are a prime option for addressing players' concerns regarding building causal relationships during self-reflection. Thus, in the fourth and final thrust of this dissertation, I explore the use of process-visualizations as retrospective gameplay visualizations and the impact that they have on self-reflection, learning, and performance.

Part V

Reflecting on and Learning through Process Visualizations

Following the work discussed in the previous part, I arrived at the conclusion that process-visualizations can provide players with a better understanding of cause and effect, improve self-reflection, and, as a result, lead to better learning and performance improvement. Building upon this argument, in this final thrust of my work, I explore the extent to which this is the case. First, I examine an open question regarding what data should be included in such presentations, specifically looking at whether or not community data should be included as a way to improve self-reflection. Finally, I test whether or not data presented in a process-oriented manner actually improves self-reflection and performance compared to data presented in an aggregate manner.

Chapter 10

The impact of One's Own and Others' Process Data on Self-Reflection

The work presented in this chapter was originally published at CHI 2023 [1]

10.1 Reflection, Adaptation, and Community

As seen from the discussion of Self-Regulated Learning, one of the most important elements of the learning process is reflection. Even outside of SRL, many learning theories formalize reflection as a central component within the learning process [590, 584]. There are even specific theories of reflection and how it occurs, including frameworks that quantify reflection across various levels [455, 680, 529, 721, 278, 32, 386, 219]. An example of such a framework is that of Leijen et al. [386] who quantify reflection across four levels (description, justification, critique, discussion), which describe *how* the student is reflecting, and three foci (technical, practical, sensitizing), which describe *what* the student is reflecting on.

[1]Kleinman, E., Villareale, J., Shergadwala, M., Teng, Z., Bryant, A., Zhu, J., & Seif El-Nasr, M. (2023). "What else can I do?" Examining the Impact of Community Data on Adaptation and Quality of Reflection in an Educational Game. Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23), April 23–28, 2023, Hamburg, Germany. ACM, New York, NY, USA, 20 pages. This research was part of an NSF-funded project conducted collaboratively with the PXL lab at Drexel University and ITU Copenhagen. It would not have been possible without the assistance of Jen, Murtuza, Zhaoqing, and Andy and the guidance and input of Jichen and Magy.

According to the theory of SRL, one key aspect of reflection is adaptation or the ability to change one's approach to a task or problem [756], and this construct is included in other frameworks as well, sometimes as a part of reflection and sometimes as a result. Using Leijen et al.'s framework as an example again, the highest level of reflection, according to the framework, is discussion, defined as "Moving beyond the evaluation and explanation of what is, and why they think that is, and pointed out what could be done to initiate changes, and why changes are needed in the first place" [386]. In other words, when one reflects at the highest level, they put cognitive effort into considering how to adapt their behavior. Many learning theories also discuss how high-quality reflection is a mechanism for change, which ultimately results in learning progress, as change is the means by which one is able to recognize and overcome mistakes and, effectively, learn.

Prompting change through reflection can, however, be exceedingly difficult, especially if the learner struggles to identify what they did wrong or what alternative strategies they could employ. This is especially the case, as illustrated in the previous chapters, in complex gaming environments where there are numerous correct solutions and many unpredictable outcomes. Within learning sciences, there is a movement that seeks to explore reflection in the context of the community, meaning reflection prompted and aided by other learners, who may be on the same level as the learner or at a higher level of expertise, and how this social reflection may improve adaptation. This movement brings to mind the theories of co and socially shared regulated learning, discussed in earlier chapters, which argue that learning incorporates the input of others, through feedback or cognitive guidance, to aid in the adaptation of one's techniques through reflection [273]. This suggests that having others present in the reflection process to help a learner perceive alternative paths can prompt and improve adaptive practices.

Having others perform this role, however, is not always a viable option. Thus, there is a need to explore alternative ways of bringing others into one's reflection process. One possibility, which I explore here, is presenting community performance data to the learner during reflection as a way to prompt adaptation. Existing work in the learning sciences has demonstrated benefits to having learners view the data of their peers while reflecting on their performance [74, 501, 291, 229]. Using community data to elicit reflection and adaptation is a focus of student-facing learning analytics dashboards (LADs), which are data-driven visual displays that "aggregate multiple visualizations of different indicators about learners, learning processes, and/or learning contexts" [73]. Student-facing LADs aim to help students make strategic decisions in learning environments related to resource (time and energy) management considerations such as what assignments to focus on, how long to study for an exam, or how often to interact with a course management system, such as Blackboard [578, 501]. In this context, being able to view the behavior of classmates and where they stand against their peers in terms of accomplishments is valuable in helping students adjust where they are spending their time and what they are prioritizing as the course progresses [578, 579, 113].

Certain gaming contexts have also adopted this approach, presenting community data through retrospective visualizations, such as those discussed earlier, to elicit reflection and motivate adaptation and more efficient learning. [80, 280, 447]. Many of the commercially available assistants discussed in earlier chapters integrate community data and comparison with that data to some degree [601, 459, 525, 68]. The benefits of learning from others in gameplay contexts, especially esports, are apparent in the work of Wallner et al., who conducted an interview and survey study examining how players use retrospective visualizations in esports [708]. Their results include themes focused entirely on what information players want or need about their opponents and illustrate how players use retrospective visualizations to learn from others. Looking at the domain of educational games, Villareale et al. [693] used existing frameworks to conduct a review of 12 programming games and identified four features used to elicit reflection. Among these is "social discourse" or "a space in the interface for community-based discussion where students can examine multiple perspectives and receive feedback on their process that can then be used for reflection" [693]. The authors go on to discuss how, through social discourse, players in programming games are exposed to different perspectives on the same problem and become more aware of alternative approaches, encouraging adaptation.

The use of community data to elicit reflection and adaptation in games is, however, under-studied, and, to my knowledge, there is no work that explicitly demonstrates that exposure to community data, especially presented in the form of a process visualization, does result in a willingness to try a different approach. Further, work on games at large has uncovered some drawbacks surrounding the use of community data. In one example, Esteves et al [196] found that social comparison in games could lead to disengagement if the player felt that they were under-performing compared to their peers. Even if disengagement does not occur, feelings of inadequacy, such as those observed by Esteves et al. [196], could result in lower quality reflection, which may impact, not only adaptation but learning as a whole. Thus, in this first study in my final research thrust, I explore the value of including community data in a retrospective process visualization, in order to provide actionable insights into whether or not the inclusion of such information is of value or a detriment to players.

10.2 A Study of the Impact of Others' Data on Adaptation

In this study, I conducted a within-subjects experiment examining the impact that comparison with peers within a retrospective process visualization had on a player's willingness to adapt one's strategy and on their quality of reflection. Specifically, this study asked the following questions:

- How does comparison with peer data impact a player's willingness to consider a different approach?
- How does comparison with peer data impact a player's quality of reflection?

I answered these questions once again using *Parallel*, described above. The insights from this work can not only help future designers make informed decisions, but also provide a flag for future research to further expand our understanding of adaptation in complex games, its relationship with visualization and reflection, and how we can prompt it.

10.2.1 Methods

For this study, *Parallel* was hosted on a web domain (playparallel.com). The game was instrumented to collect and save information about all of a player's actions during level seven in log files. Level seven was chosen for this study because it was a complex enough level to be a reasonable challenge to participants and could be solved in several ways of varying correctness. As such, it warranted some manner of reflection and adaptation, but not so much of a challenge that participants may become overwhelmed or fail to complete the level.

10.2.1.1 Recruitment

36 undergraduate computer science students, the intended user group for *Parallel*, were recruited from programs at UCSC, Northeastern, and Drexel University. Participants were required to be 18 years of age or older, located in the United States, able to communicate in written and spoken English, and able to access the playparallel website, but were not required to have prior experience with parallel programming. That being said, 17 did.

10.2.1.2 Protocol

Upon giving informed consent, participants proceeded through the following steps:

- **Account Creation and Tutorial:** Participants were provided with instructions for how to access playparallel.com, create an account, and complete the *Parallel* tutorial. They were given five days to complete this step.
- **Level 7 (First Playthrough):** The day after the deadline to complete the tutorial, participants were sent instructions to play level 7 of *Parallel*. They were

given three days to complete this step.

- **Level 7 (Second Playthrough):** The day after the deadline to play Level 7 the first time, participants were asked to play level 7 a second time. They were given three days to complete this step.

- **Reflection 1:** The day after the deadline to play Level 7 the second time, participants were randomly assigned to a reflection condition and provided with a visualization setup, either self or peer (see below), and responded to a set of reflection prompts (see below). They were given three days to complete this step.

- **Reflection 2:** The day after the deadline to complete the first reflection step, participants were provided with a second visualization setup for whichever reflection condition they did not do the first time (see below), and responded to a set of reflection prompts (see below). They were given three days to complete this step.

This protocol was run remotely and asynchronously. Participants were compensated twice during the protocol. They received 20$ after completing *Reflection 1* and 30$ after completing *Reflection 2*. University IRB reviewed and approved the protocol.

10.2.1.3 Visualization Setup

For the reflection steps, I created a process-visualization for each player using a variation of the network graph from *Glyph* [478] to visualize players' gameplay as a sequence of actions. This is the same system that was used in the previous study with *Parallel*, and the same set of key actions was used to generate the visualizations in this study.

In addition to the action name, for this study, each node, in the node title, also indicated where, on the game board, the respective action was taken. To facilitate this, the game board was split into sections based on the shape of the track and the location of pick-up and drop-off spots, as seen in Figure 10.2. Based on this game board abstraction, each node in the process-visualization would contain, after the action name,

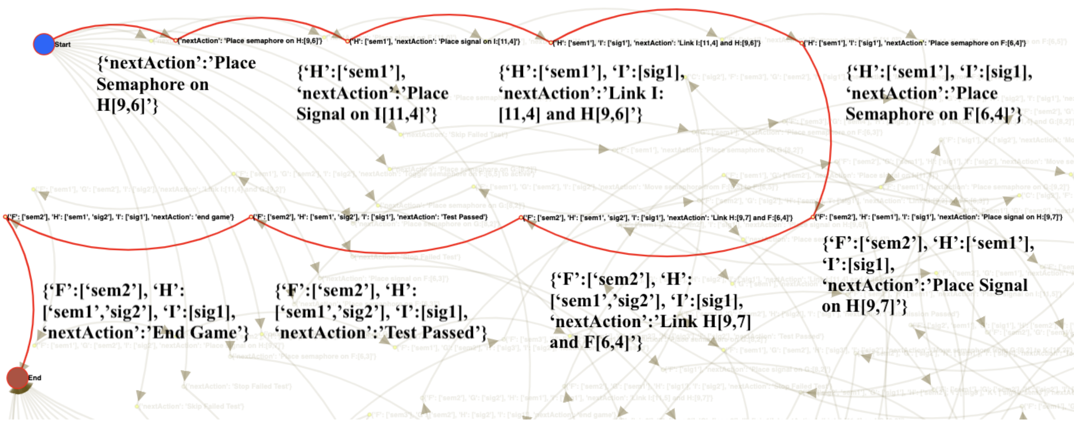


Figure 10.1: A play trace depicted in *Glyph*'s network graph. For readability, I have enlarged the text in the labels.

an indicator of which section of the board the action occurred in along with the exact coordinates, which could be used to differentiate between the same elements in the same sections (i.e. two semaphores in section g could be differentiated by their coordinates). For example, a node may say "place semaphore on H:[9,7]" to indicate that that action was the placing of a semaphore at the coordinates 9,7 in board section H. Additionally, each node label would also display which board sections had elements in them already. For example, a node may say "F:[sem2,sig1]" indicating that the second semaphore placed in this playthrough and first signal placed in this playthrough are in section F. How these labels looked in the visualization can be seen in Figure 10.1. To better convey location and board state information, each visualization was augmented with images of the board state at every action taken, as seen in Figure 10.3. Visualizations also included the key image seen in Figure 10.2 so that players knew what the sections of the board were. I created each visualization manually by combining screenshots of the *Glyph* output with the board images in *Miro*.

10.2.1.4 Reflection Prompts

The repeated measures study featured two reflection steps, which I refer to as "self" and "peer" reflection. During **self-reflection**, a participant was shown a

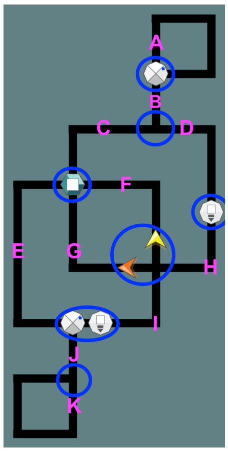


Figure 10.2: The division of the level 7 board into sections.

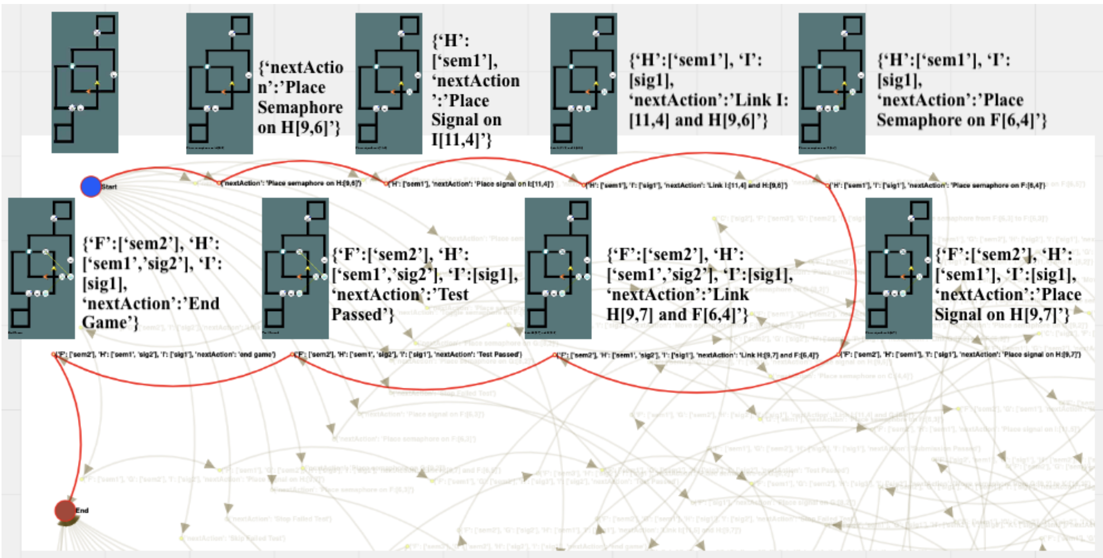


Figure 10.3: A playtrace with both the node link diagram and the images depicting the board state at every action. For readability, I have enlarged the text in the labels.

visualization of their own gameplay for their first playthrough of level 7 alongside their own gameplay for their second playthrough. An example of this visualization setup can be seen in Figure 10.4. During **peer-reflection**, a participant was shown a visualization of their own gameplay for their first playthrough of level 7 alongside two other players' gameplay for level 7, one that was similar to theirs and one that was different. The similarity was determined using *Glyph*'s sequence graph feature [478]. Participants were told which trace was similar or different. The other players' traces could have been a first or second playthrough, but participants were not informed of this. An example of this visualization setup can be seen in Figure 10.5.

During a reflection step, a participant was directed to their respective Miro board and provided with a short video on how to interpret the visualization. They were then asked to respond to a set of questions in a google form. For the peer reflection step, these were as follows:

- Please look at your gameplay sequence for your first attempt. Based on your sequence, can you describe how you approached the level?
- Please look at P1's gameplay sequence. Based on their sequence, can you describe how they approached the level?
- Please look at P2's gameplay sequence. Based on their sequence, can you describe how they approached the level?
- Compared to the other players, what went well in your playthrough and why?
- Compared to the other players, what went poorly in your playthrough and why?
- If you were to play this level again, would you do anything differently?

For the self-reflection step, these were as follows:

- Please look at your gameplay sequence for your first attempt. Based on your sequence, can you describe how you approached the level?

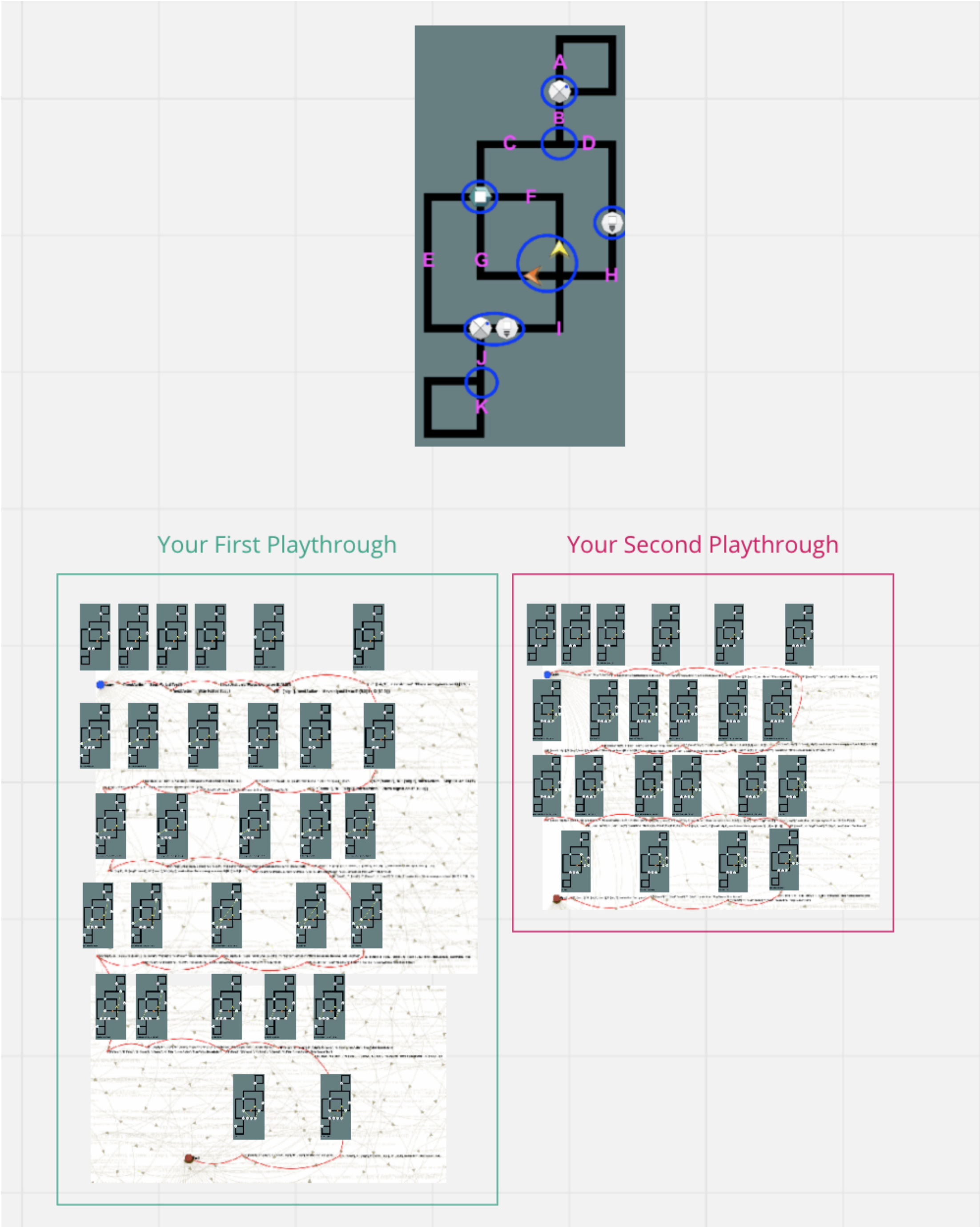


Figure 10.4: The visualization setup for the self reflection condition depicted the playtrace for the player's first and second playthroughs of level 7.

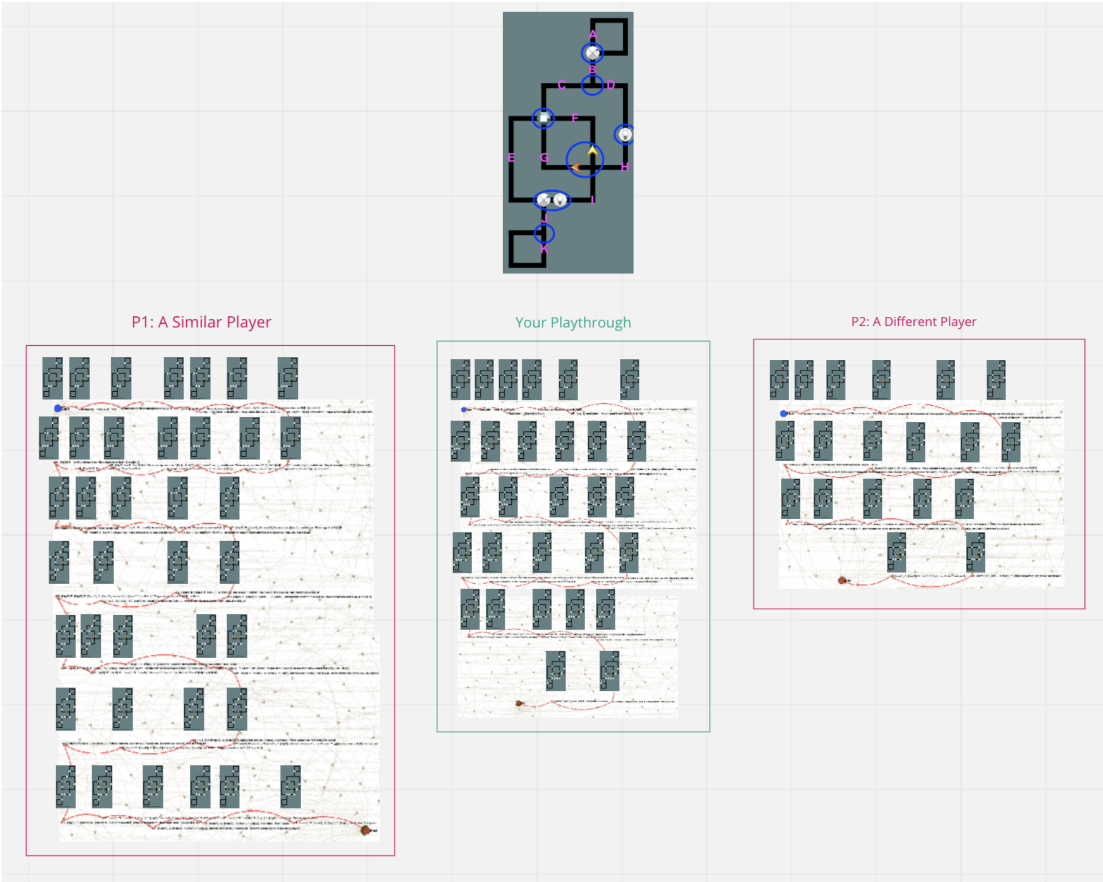


Figure 10.5: The visualization setup for the peer reflection condition depicted the playtrace for the player's first playthrough of level 7 alongside two other playtraces, one similar and one different.

- Please look at your gameplay sequence for your second attempt. Based on your sequence, can you describe how you approached the level?
- Compared to your second attempt, what went well in your first playthrough and why?
- Compared to your second attempt, what went poorly in your first playthrough and why?
- If you were to play this level again, would you do anything differently?

In both conditions, only the last three questions were meant to elicit reflection, with only the very last question being related to adaptation. The first batch of questions in both conditions existed to ensure that the participant looked at all the playtraces in enough detail to be able to reflect on them. These reflection questions were derived based on similar questions and prompts used in previous work [364, 454, 58, 186]. The order of reflection was randomized with half the group doing peer-reflection first and the other half doing self-reflection first, but all participants reflected in both conditions. The last question was answered with a yes or no while all others were open-response.

10.2.1.5 Data Analysis

Open-answer responses to the “what went well?” and “what went poorly?” prompts were analyzed qualitatively using Leijen et al.’s reflection model to measure the quality of reflection [386]. As discussed earlier, the model quantifies reflection based on focus and level as seen in Table 10.1. I worked with a collaborator to define what these concepts mean within the context of *Parallel* and concluded on the definitions seen in Table 10.2.

I then worked with the same collaborator to, separately, apply the codes to half of the data set. The unit of analysis was a single response and one code for focus and one code for level was applied to each unit. Inter-rater reliability was measured

Label	***Definition***
Focus	
Technical	Concerned with the efficiency of means for reaching certain goals
Practical	Involves an open examination, not only of means but also of goals, the assumptions goals are based on and the actual outcomes
Sensitizing	Concerned with social, moral, ethical, or political aspects
Level	
Description	Mere descriptions of actions and thoughts
Justification	A rationale or logic for an action or viewpoint
Critique	An evaluation for an aspect and explained why this explanation was given
Discussion	Moving beyond the evaluation and explanation of what is, and why they think that is, and pointed out what could be done to initiate changes, and why changes are needed in the first place

Table 10.1: Leijen et al.'s [386] model for measuring the quality of reflections.

Label	***Definition***
Focus	
Technical	Discussing efficiency in terms of what the player did, does not include discussion of goals, but may include statement of a goal.
Practical	Discussion of goals, what they were, how they changed, if they were good or bad, etc...
Sensitizing	Discussion of more than just goals and actions taken, thoughts about the player's status as a learner or a player, etc...
Level	
Description	Simply describing what the player did or were thinking
Justification	Providing some kind of explanation or defense or justification for why the player did what they did
Critique	Discussions of how well the player did or any kind of evaluation of their process
Discussion	Any discussion of doing things differently, next steps, what would be done if the step was repeated or done differently next time

Table 10.2: The definitions I derived for how Leijen et al.'s model applies in the context of Parallel.

Condition	***Yes***	***No***
Would you do anything differently?		
Self	2	34
Peer	12	24

Table 10.3: Differences in willingness to do something different next time between self and peer reflection.

using Cohen's kappa [133] and an agreement of .85 was reached for focus and .88 for level, both indicating very strong agreement [375]. I then coded the entire dataset.

Once the codes were applied, McNemar-Bowker tests [368] were used to determine if there were any significant differences in how often each level and focus code for reflection was applied to responses in each condition (peer and self). McNemar-Bowker tests were also applied to the "yes/no" responses to the last question, which was used to determine if exposure to community data and reflection through comparison with that data impacted willingness to consider a different approach.

10.2.2 Results

The counts for how many participants said they would consider a different approach next time after completing each reflection step can be seen in table 10.3. These results indicate a notable increase in the number of participants who indicated that they would try something else if they played again after reflecting on peer data.

McNemar-Bowker tests revealed that this difference was significant (p=.004), with players being significantly more willing to consider an alternate approach (adapt) when reflecting on their own data in the context of others'. Effect size, calculated using Cramer's V, resulted in an effect size value of .5, indicating a large effect.

The counts for the focus and level codes in both conditions can be seen in Tables 10.4 and 10.5, respectively. Technical reflection was most common by a large margin in both conditions and sensitizing reflection was the least common. Similarly, description and justification were far more common across conditions than critique or discussion. The tables also indicate that there was a slight increase in higher quality

Condition	***Technical***	***Practical***	***Sensitizing***
What went well?			
Self	27	5	4
Peer	29	5	2
What went poorly?			
Self	28	7	1
Peer	24	9	3

Table 10.4: Differences in focus of reflection between self and peer.

Condition	***Description***	***Justification***	***Critique***	***Discussion***
What went well?				
Self	20	11	4	1
Peer	21	10	3	2
What went poorly?				
Self	18	12	4	2
Peer	16	11	2	7

Table 10.5: Differences in level of reflection between self and peer.

reflections (those of the sensitizing focus or discussion level) during peer reflection when players were asked to reflect on what went poorly. McNemar-Bowker tests, however, revealed no significant changes between conditions (all p ¿ .05).

10.3 Considerations for Using Community Data to Influence Self-Reflection and Adaptation

While previous work in the learning sciences has demonstrated the value of reflecting on community data as a way to encourage adaptation in learning [578, 294], community data and retrospective visualizations are under-explored in educational games. As a result, the impact of community data on adaptation within the domain has not been empirically examined and, further, previous work in other game genres has suggested that exposing players to peer data may have a negative impact [196]. Based on this, my goal was to determine whether or not comparison with community data, specifically presented via process visualization, in a complex game had a significant impact on players' willingness to consider alternative approaches and, alongside this, whether or not it had a significant impact on the quality of reflection. I pursued this goal in order to provide researchers with foundational knowledge upon which additional research may be conducted and to arm developers with insights that would allow them to make informed decisions with regard to when to leverage peer data.

The results indicate that reflection on one's own data in the context of peers' data does, at least partially, motivate adaptation, suggesting that including community data in process-visualizations will have a positive impact as far as supporting SRL. This is highlighted by 1/3rd of the participants in this study demonstrating a significantly higher willingness to consider a different approach if they were to play the level again after peer reflection. This is an important finding, as previous work in Learning Analytics Dashboards has alluded to a willingness to try a different approach as a benefit of community data [229, 579, 501] but has not empirically explored the topic. This has also not been explored extensively, to date, in the context of complex games, although it has been alluded to [80]. It is further valuable to identify this relationship in the context of the concerns surrounding the use of community data in games, which could lead to discouragement and disengagement [196]. Effectively, these results provide empirical evidence that there is a benefit to including community data in process-visualizations meant for self-reflection in complex games. It can help players, especially those who may be struggling, entertain alternative approaches they may not otherwise consider or perceive and more efficiently arrive at a correct solution, meaning they can move on more quickly and continue to advance and improve.

Notably, however, only about 1/3rd of participants said they would do things differently after reflecting in the peer condition. This suggests that, while the presence of community data has a significant effect, it is not an all-encompassing way to prompt adaptation. It is likely that there are other ways to prompt adaptation both in the context of retrospective visualization and as a part of reflection in general that may result in a greater number of people deciding to try a different approach. Further, there may be individual characteristics, such as confidence or stubbornness, that may impact an individual's willingness to adapt based on comparison to peer data alone, similar to how competitive preferences impact self-monitoring and learning in Riemer and Schrader's work [558]. As such, these results should not be taken to mean that including community data is the single, best way to prompt adaptation in complex games.

Instead, I present these results as a foundational understanding of the impact of peer data on adaptation. Specifically, I suggest that peer data encourages adaptation, but not for everyone, and may be one of many possible techniques for prompting change among learners, and may be impacted by individual characteristics. **As such, I suggest that developers of retrospective visualizations for complex games consider leveraging comparison with peer data as a way to motivate adaptation, but remain open to alternative techniques as well, as peer comparison alone will not prompt the entire population**. Additionally, **I recognize an opportunity for future work to explore the alternative ways that adaptation can be prompted, either as an alternative to peer comparison, or in combination with it, and the ways in which player characteristics influence willingness to change.**

Given these implications, it is important to note that I also did not observe a significant change in the focus or level of reflection when peer data was introduced. Reflection for both groups favored the technical focus and the description and justification levels, which reflects the findings of previous work that found that reflection in games often does not rise to the highest levels [453]. This, however, does not conclusively mean that there is no effect of peer data on the quality of reflection. While I controlled for individual differences through a within-subjects design, it is possible that a larger sample size or different context could lead to a different outcome or significant change. Given the inherent risks of peer data impairing reflection or learning quality, such as by prompting a player to merely copy what they saw without thinking about it or, as demonstrated by previous work [196], to completely disengage with a game, further exploration of this question remains relevant. As such, **I recognize these findings as**

motivation for the further exploration of the topic in future work.

These results also raise additional questions about when to expose the community data in order to elicit change. Schon [584] describes two types of reflection in terms of when they occur: reflection *in action* and *on action*, with the former referring to reflection occurring during an event and the latter occurring after the fact. In this study, I specifically examined reflection on-action, as my interest is in retrospective process-visualizations and how the presentation of process information impacts self-reflection. It may be that reflection in action, which would be a performance phase activity within the structure of SRL, produces different results. For example, more participants may be willing to consider an alternative strategy if they have not yet completed the task. It is also possible that the quality of reflection could be impacted by peer data when the reflection occurs in the midst of the activity. Such a suggestion also aligns with LAD work that found that students liked seeing other students' approaches so that they could evaluate their standing and adjust their approaches before completion of a course [579]. Reflection in action, however, involves short cycles of thinking and doing, and there may not be enough time in such a structure for a player in a complex game, especially a fast-paced one, to make meaning of community data, which, as the previous chapters illustrated, is a complicated cognitive process. This drawback could potentially prompt the above-mentioned negative behavior of copying without thinking. As such, **I recognize an opportunity for future work to explore the open question of when, in relation to the progress and completion of a task within an educational game, reflection on community data should be presented to, not only elicit change but, lead to success.**

In relation to these considerations, previous work has also found that exposing users to the best approach, in situations where there is a single best approach, may result in conformity among a population [326]. In other words, making the players of any given complex game aware of what the best solution is, if one exists, could result in all players following that same solution. In some circumstances, this may be ideal, such as in the contexts of Learning Analytics Dashboards or Open Learner Models where students are encouraged to “follow in the footprints” of other more successful students [300] or when the goal of the game is to help students arrive at and understand a single correct solution. In circumstances where there are multiple correct pathways, or no “right” answer, which is often the case in complex gaming environments, as I discussed earlier, however, sharing community data that could lead to conformity may not benefit the players as much as it could reduce creativity or variety. This illustrates open questions regarding how to best present community data to players of a complex game, especially in contexts where there is not a single correct answer or conformity is otherwise not desired. Thus, **I recognize an opportunity for future work to explore how to present community data such that the “best” or “correct” solution is not exposed in such a way that inhibits players’ ability to explore and learn naturally. Further, I see opportunities for future work to explore how this presentation should differ depending on the academic context and design of the game.**

Finally, I acknowledge that there exist a number of additional open problems surrounding the fair use of community data. For example, Zhu and El-Nasr [746] discuss how public player data raises concerns about privacy and ethics in open player models. These concerns resonate with considerations that players brought up in the studies discussed in the earlier chapters of this thesis, suggesting that publicly available data could result in toxic behavior or foul play. The findings from these earlier studies mirror what was found by Park et al. when they conducted a requirements analysis before designing their learning analytics dashboard [501], suggesting generalizable concerns. While the work presented in this chapter does not address these questions of privacy, fair use, and fair play, I argue that, by demonstrating the value of community data within the domain of educational games, this work motivates the exploration, and hopefully resolution, of these open problems. Thus, **I recognize an opportunity for future work to use these results as motivation to explore open problems surrounding the social and safety concerns inherent in the use of community data.**

10.4 Summary

In this chapter, I discussed a study that explored the value of including community or peer data in a post-play process visualization meant to support self-reflection. While I hypothesized that the inclusion of community data would prompt adaptation, a key element of self-reflection, and found that it did do so, to an extent, the results ultimately suggested that this was not a single, sure-fire way to promote adaptation or improve SRL. Given these results, I move forward in this thesis work to my final study, in which I explore the impact of a process visualization on self-reflection when compared to an aggregate visualization. Based on the results discussed in this chapter, I further choose to explore this question by looking only at visualizations of the player's own data.

Chapter 11

The impact of Process Data on Performance in Complex Gameplay

The work presented in this chapter is currently under review at CHIPlay 2023 [1].

11.1 Reflection, Performance, and Visualization for Games

While the previous study explored the impact of community data on adaptation, it did not look at the bigger picture of whether or not a process-visualization, in and of itself, had an impact on SRL and performance in a complex game. As such, this final study in this thesis sought to explore this question and provide foundational and valuable insights into the connection between the presentation of gameplay data, SRL, reflection, and performance in complex games.

As discussed earlier, previous work has barely investigated the impact of visualized data on learning in games, instead focusing predominantly on design and usability [274], despite using learning as a motivator for the creation of new tools [701, 11, 367]. Similarly, previous work on reflection in games has not examined its impact on learning gameplay itself, instead, focusing on reflection on narrative experiences [453, 119] or an

[1] I led this research but it would not have been possible without the assistance of Jason Xu and the guidance of my advisor Magy.

educational topic [72, 569]. In addition to this, little work has explored the connection between reflection and data visualization in games, with the previously discussed work of Villareale et al. on reflection prompts in educational games [693] being one of the only examples. But even here, the types of visualizations they examine in the games they review vary dramatically from process-visualizations as they are defined within the context of this thesis.

Beyond games, there has been a more extensive investigation of the impact of data on learning and reflection. Student-facing Learning analytics dashboards [73] (LADs) have been used to help students use data visualizations to compare their performance against their peers [501] and track progress towards learning outcomes [113]. User study results have found that data visualized through an LAD can support students' reflection and strategy planning [138], help them get a better idea of how their activity with course material may impact their performance [229], and help them make more informed decisions about how they spent their resources [579].

While LAD research has not looked extensively at reflection, as a process and as a part of SRL, Open Learner Models (OLMs), which are often informed by SRL, have. OLM research has demonstrated, through user studies, that presenting learners with their model, which is a representation of their performance data, can help them better identify gaps in their knowledge [276, 380, 121]. OLMs, however, do not often support reflection through basic data visualization. Instead, as discussed in Chapter 2, many have taken the negotiation approach, in which the student is prompted to debate the accuracy of their model with the system, a process that prompts reflection [647, 162].

In summary, the connection between data visualization, reflection, and learning is under-explored in complex games, making it difficult to know how to design computational tools to support SRL and improve learning and performance within the domain. While explored more extensively in learning sciences, the findings of this literature are not necessarily applicable in the context of learning gameplay. This is because game data is inherently more complex, resulting in unique interpretation techniques and user requirements, as demonstrated by the results of the previous thrust. Further, previous work, in both games and learning sciences, examines learning and reflection in the context of academic content, not gameplay performance itself. Additionally, none of the previous work looks at process-visualizations as I define them in the context of this dissertation. As such, in this final study, I propose to explore, specifically, the impact of process-visualizations of one's own gameplay data on reflection and learning.

11.2 The Impact of Process on Reflection and Performance: A User Study

To address this gap, I conducted a between-subjects user study examining whether or not reflection on a process visualization, versus an aggregate one, had an impact on reflection or performance in the context of League of Legends. Specifically, this study asked the following two questions:

- How does a process visualization impact performance in an esport game compared to an aggregate visualization?
- How does a process visualization impact self-reflection on gameplay in an esport game compared to an aggregate visualization?

I once again returned to the domain of esports for this final study. As stated previously, esports offer one of the most complicated complex gaming experiences. Further, the results of the studies with *Parallel* revealed to me that players will often reflect harder and better when engaging with data from a game they are familiar with and passionate about.

11.2.1 Methods

11.2.1.1 Recruitment

28 intermediate (self-identified) League of Legends (LoL) players were recruited through an online community, word of mouth, and snowball sampling. Recruitment

information was initially posted in the Discord server for *Mobalytics* [459] and was spread by individual participants through their personal social circles. Similar to the earlier LoL study discussed in part 1 of this thesis, intermediate refers to someone who currently played (or had played in the past) in a recreational, and occasionally competitive manner, either solo or on a team, but did not, and had not, regularly play at the competitive level. Demographic information was collected on the recruitment form after the participant gave informed consent.

11.2.1.2 Gameplay Task

Like most esports, LoL is a large, dynamic, multiplayer game, meaning that the normal online gameplay mode contains numerous external variables that cannot be controlled. As such, to accurately measure changes in performance, participants were asked to complete a specific exercise in the game's training mode. LoL's training mode allows players to create custom games with as many human or AI opponents as they would like. For the purposes of this study, the player would be on the green team, and an AI of the champion Zyra, set at intermediate difficulty, would be on the red team. I chose Zyra for the opponent after reviewing community materials and determining that she was generally considered to be the most difficult AI to play against. This meant that she posed a substantial enough challenge for the purposes of the task. To ensure that familiarity with a champion would not be a confounding variable, participants were allowed to select any champion they were comfortable with for the task. They were also allowed to use any summoner spells, special abilities that are not tied to a specific champion, and runes, a kind of stats page, they desired.

The gameplay task gave participants 10 minutes to earn as much gold as they could. They were allowed to do anything they wanted to achieve this goal, though, technically speaking, there are only four ways to earn gold within the confines of the task: kill creeps, kill creatures in the jungle, kill the enemy champion, and destroy the enemy towers. Players were allowed to return to their base and shop whenever they wanted. The only requirement was that players had to begin the task in the bottom lane, as that was the lane that the Zyra AI would always begin in. Participants were additionally informed that they would be deducted 100 gold from their final score for every death or lost tower.

The task was designed to permit variety in what players could do to complete it. They could focus on farming, killing Zyra, killing jungle NPCs, taking towers, etc... It also afforded them the opportunity to change their strategy when they repeated the task, which I discuss further in the next section. A performance score for each player was calculated based on the total amount of gold earned by the player, minus the deductions for deaths or lost towers. However, no participant ever lost a tower, so points were only ever deducted for deaths.

11.2.1.3 Protocol

Upon giving informed consent, each participant was randomly assigned to either the “process” or “aggregate” group, which would determine how their data would be visualized. I then contacted them to schedule two zoom sessions, one day apart. In the first session, I asked the player to complete the task while I recorded the screen and audio. The player was given the opportunity before, during, and after the task to ask any questions. This first session ranged from 15 to 30 minutes and participants were not required to think aloud during gameplay. At the end of the session, I recorded the total amount of gold earned, the number of deaths, towers lost, creeps killed, and enemy champion kills. I then calculated the performance score by subtracting 100 gold for each death (no towers were ever lost) from the total amount of gold earned.

After the completion of the first session, I reviewed the gameplay recording and, leveraging my approximately 10 years of experience with LoL, labeled the actions taken by the player using the action labels presented in table 11.1. These labels are adapted from the coding scheme presented by Ahmad et al. [14]. The result of the labeling process was a sequence of actions taken by the player. This would then be converted into either a process visualization or an aggregate visualization, depending on the player’s group, as seen in Figures 11.1 and 11.2. More details on how the

Action	Definition
Farm	An uninterrupted period of actively last-hitting minions to obtain gold
Harass	An uninterrupted period of actively attacking the enemy champion
Return	Teleport back to the base using the return hot-key
Shop	Purchase or sell items at the base shop
Roam [Location]	Move across the map to the specified destination
Teleport [Location]	Use the teleport summoners spell to move to the specified destination
Jungle	An uninterrupted period of attacking and Killing jungle creatures
Dragon	Killing the dragon
Push Tower	An uninterrupted period of actively hitting a tower
Push Lane	An uninterrupted period of actively hitting, but not last hitting, creeps
Kill	Kill the enemy champion
Die	Get killed by the enemy champion
Destroy Tower	Destroy an enemy tower
Lose Tower	Lose an allied tower
Turtle	An uninterrupted period of positioning oneself within the firing range of an allied tower
Level [Skill]	Level up the specified skill

Table 11.1: 16 possible actions that players could take during the gold collection task, modified from the behavioral abstraction presented by Ahmad et al. [14].

visualizations were set up are provided below.

In the second session, which would occur one day after the first, I would first conduct a Data-Driven Retrospective Interview [186], where I showed the participant the visualization of their gameplay data and asked them three questions meant to prompt self-reflective processes. The questions were adapted from Cleary et al.'s study of Self-Regulated Learning in basketball players [132] and each prompted one of the three main processes of self-reflection according to Zimmerman's model [756]: evaluation (how learners decide if they did well or not), attribution (identification of the causes of failures), and adaptation (identification of what to change):

- Please evaluate your performance during this game and elaborate on how you are reaching your conclusions (Evaluation)

- Please identify something during this game that you think you did not do particularly well, please elaborate on what you think caused your poor performance (Attribution)

- What do you think you need to do to perform better in your next game (Adaptation)

The zoom call for the second session was also recorded to capture participant responses.

Immediately following this interview, the participant repeated the task from the first session. They were required to use the same champion, runes, and summoner spells as the first session (they were informed of this when they picked these for the first session), but once they proceeded to the bottom lane within the game itself they were allowed to change what they did.

Following the completion of the second gameplay task, I asked each participant four additional questions. With the exception of the third question, these questions were adapted from Kitsantas and Zimmerman [343] and my earlier study with LoL players from part 1 of this dissertation [345]. The questions examined three processes related to Self-Regulated Learning: Goal Setting, Strategy Use, and Self-Monitoring. While not explicitly a part of self-reflection, these processes are still critical elements of SRL under the CPM model, and the goal of this section of the protocol was to examine if the style of visualization had an impact on any other parts of SRL beyond self-reflection. The questions were as follows:

- Did you set any specific goals for the match and if yes, what were they and what prompted them? (Goal Setting)

- What did you need to do to accomplish your goals? (Strategy Use)

- Were you able to do so? (Goal Realization)

- Did you monitor your progress during gameplay? (Self-Monitoring)

Following the completion of the second set of interview questions, participants were given the chance to provide any additional input and ask any questions. The second session took between 30 - 45 minutes. Participants received a 25$ gift card for completing each session. The two sessions were scheduled one day apart to facilitate the creation of the process-visualizations for each player and to ensure enough time had passed that the player would not be relying entirely on their memory of their gameplay when reflecting, encouraging them to interact with the visualization, but not so much that they could not remember anything. University IRB approved the protocol.

11.2.1.4 Visualization Setup

I manually created visualizations for players in both groups using Miro, in order to create something that players could easily access, understand, and interact with on their own computers, regardless of technical specifications or prior experience with visualization. Examples of visualizations for both groups can be seen in Figures 11.1 and 11.2. For actions such as "farm", "harass", or "jungle", each node indicated an uninterrupted period of taking that action, as opposed to a single strike. Action nodes were assigned a unique color for readability, the colors were the same for both groups and both groups were provided with a key for what each action was, seen on the right side of each visualization in both figures.

For the process group, the visualization displayed a timeline-like sequence of the actions they took, in the order they took them, seen in Figure 11.1. For readability, this timeline would double back on itself in a zig-zag manner so that the entire sequence could be seen without needing to scroll. The point at which it doubled back had no significance and participants were informed as such. For the aggregate group, the visualization displayed the aggregate count for how many times they took each action, seen in Figure 11.2.

Existing in-game and third-party visualizations typically aggregate data such

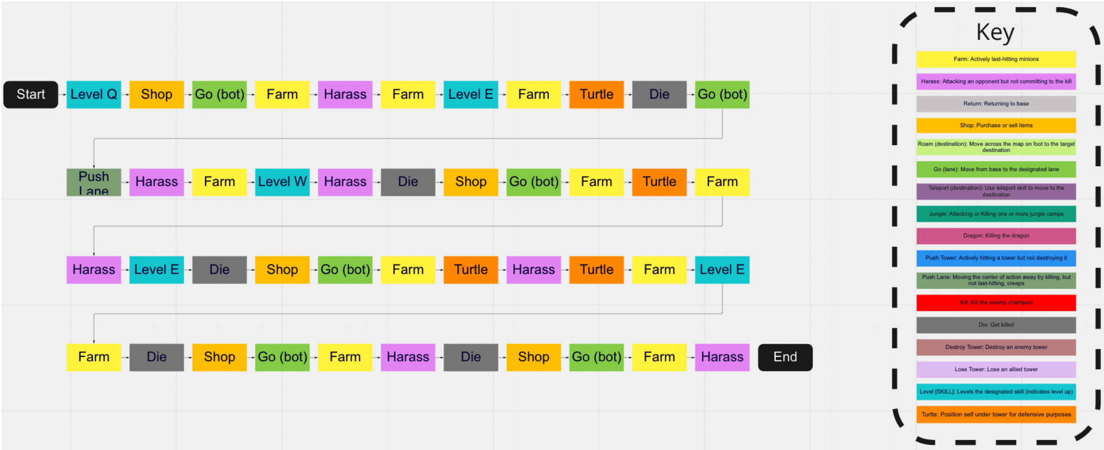


Figure 11.1: An example of the process visualization with the key on the right.

Figure 11.2: An example of the aggregate visualization with the key on the right.

as gold earned, damage dealt, and the number of kills or deaths [601, 459, 68, 525]. However, to ensure that the information available to both groups was as similar as possible, the aggregate group here was shown the number of times they took each action, as the same information could also be gleaned from the process-visualization. This meant that the only difference between the two visualizations was that the process group could see the order of actions, such that I could determine if this information impacted self-reflection. Participants were provided with explanations for how to read these visualizations and how the nodes should be understood and given the opportunity to ask questions before the interview began.

11.2.1.5 Data Analysis

Quantitative Analysis To answer RQ1 "How does a process visualization impact performance in an esport game compared to an aggregate visualization?" statistical tests were run to check for significant differences in gameplay metrics from before to after reflection. Specifically:

- overall performance score
- number of kills
- number of creeps killed (Creep Score)
- number of towers taken
- number of deaths

To answer RQ2 "How does a process visualization impact self-reflection on gameplay in an esport game compared to an aggregate visualization?" various qualitative analyses, described below, were conducted on the players' responses to the interview questions. Additionally, the amount of behavioral change from before to after reflection for each group was analyzed using the method described below, in order to determine the extent to which adaptation did or did not occur.

Shapiro-Wilk tests were used to check for normal distributions of the numerical performance data and parametric or non-parametric tests were used to test for significance based on the results as follows:

- Paired T-Tests were used on performance scores and the number of kills to check for significant changes from before to after reflection.
- Wilcoxon Signed Rank Tests were used on the number of creeps killed, number of towers taken, and number of deaths to check for significant changes in performance from before to after reflection. Sign tests were used for any data for which the Wilcoxon Signed Rank test could not produce a value.
- A Wilcoxon Rank Sum Test was used to check for significant differences in the amount of behavioral change from before to after reflection (discussed below) between groups.

The Dynamic Time Warping (DTW) Algorithm presented in [348] was used to determine how much players in both groups changed their behavior between their two attempts at the task. DTW calculates differences between sequences of actions as follows: "A distance value d is created for every sequence pair (S_1,S_2). At every given step s_a in sequence S_1 the value at step s_a is compared to the value at every given step s_b in sequence S_2 and a weight w is calculated and added to d." [348]. For the purposes of this work, w was determined as follows:

- If s_a and s_b were the same action in the same location (top, mid, bot lane, or jungle) then $w = 0$
- If s_a and s_b were different actions in the same location (top, mid, bot lane, or jungle) then $w = 1$
- If s_a and s_b were the same action in different locations (top, mid, bot lane, or jungle) then $w = 2$

- If s_a and s_b were different actions in different locations (top, mid, bot lane, or jungle) then $w = 3$

In order to run this calculation, after each attempt at the task by each player was converted into a sequence of actions as discussed above I applied a location tag (indicating if the action was taken in the bottom, middle, or top lane or the jungle) to each action. Actions taken in the base (shop, return) were not tagged. The "river" location tag was only applied to the "roam" and "go" actions, as these were the only actions ever taken in relation to the river.

Qualitative Analysis The qualitative data collected while reflecting on the visualization was coded using the coding scheme used by Cleary et al. [132] that examined the self-reflective constructs of evaluation, attribution, and adaptation. For the purposes of this study, the definitions for the codes were adapted by myself and a collaborator who was also an experienced League of Legends player. For the SRL process of evaluation, codes capture what the player based their evaluation on, and were defined for this work as follows:

- **The performance of others:** They base their evaluation on comparison to community standards such as how much CS you should have per minute, how much gold you should have at ten minutes, when you should go back to shop, any type of community or meta level ideas of how the game should be played or assumptions of how other participants did.
- **Their scores:** They base their evaluation on the number of times they took an action or they reference their in-game scores (kda/cs/gold/etc...)
- **Their use of the correct method or strategy:** They base their evaluation on the execution of skills, such as timing or ordering of skills, how use of skills related to their goals, etc... May reference scores, but goes beyond just counts.
- **Their improvement during gameplay:** They base their evaluation on perceptions of improvement over the course of the task i.e. "It was rough at the start

but then things got better"

- **Other factors:** Anything else they discuss not covered
- **They don't know:** They did not elaborate on how they're evaluating themselves

For the SRL processes of attribution and adaptation, the codes captured what players considered to be the source of their failure and the element they needed to improve, respectively. The same codes were used for both processes, however, for adaptation, the "confidence/ability" code was not applied, as discussed by Cleary et al. [132]. The codes were defined for this work as follows:

- **Specific technique:** Discussions of failure to execute or need to change/improve specific strategic or technical maneuvers
- **General technique:** Discussions of failure to execute or need to change/improve general strategies or techniques
- **Confidence/Ability:** Discussions of lack of confidence or ability to perform the task (only applied to attribution)
- **Focus/Concentration:** Discussions of inability to or need to change/improve ability to remain focused, concentrate on the task, or prioritize objectives
- **Effort:** Discussions of not trying hard enough or needing to try harder
- **Practice:** Discussions of lack of practice or familiarity or needing to practice more
- **Rhythm:** Discussions of issues with or need to change/improve patience or timing of execution
- **Distractions:** Discussions of external stimuli that interfered or needing to ignore these
- **Other:** Anything not covered by the other codes

- **They don't know:** If they do not have an answer

The qualitative data collected during the final interview was coded using the coding scheme used in the earlier LoL study from part 1 of this dissertation [345], which was, in turn, adapted from Kitsantas and Zimmerman [343]. These codes examine the SRL processes of goal setting, strategy use, and self-monitoring. For goal setting, the codes were defined for this work as follows:

- **Outcome goals:** Statements related to reaching objectives or checkpoints.
- **Process goals:** Statements related to strategic gameplay maneuvers.
- **Other:** Statements that did not discuss either of the above.

For strategy use, the codes were defined for this work as follows:

- **Specific technique:** Discussions of technical execution such as using the right skill, but in relation to them actually doing it, not just thinking about doing it
- **Visualization strategies:** Discussions of visualizing or imagining oneself doing something
- **Concentration strategies:** Discussions of focusing or concentrating either in general, on a specific aspect of gameplay, or on an objective.
- **Technique and concentration:** Responses that included both.
- **Practice/no strategy:** Answers that just discussed practicing or did not discuss any strategy.

For self-monitoring, the codes were defined for this work as follows:

- **Score Alone:** Discussions of tracking the number of last hits, amount of gold, KDA, etc. either in one's head or using the in-game score board
- **Use of technique or form and its outcomes:** Discussions of technical execution of a skill or strategy

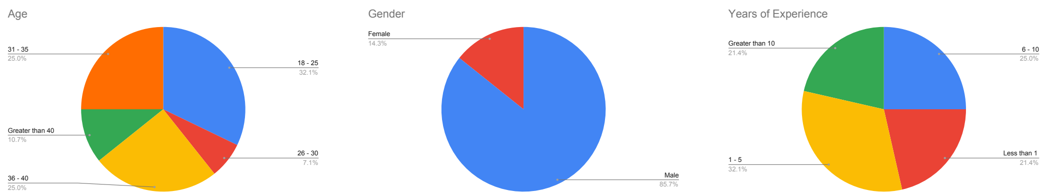
SRL Process	Resulting IRR
Evaluation	.72
Attribution	.80
Adaptation	.78
Goal Setting	.86
Strategy Use	.74
Self Monitoring	.85
Goal Realization	.81

Table 11.2: Inter-rater reliability measures for coding for all SRL processes. Calculated using Cohen's Kappa [133]

- **Do not know:** Statements indicating that they did not monitor their performance or were not sure if they did.
- **Other:** Any self monitoring strategy that did not correspond with the above.

The goal realization responses were coded with "yes", "no", and "partially". The unit of analysis for coding was an entire response. All codes were only applied to the responses to their corresponding questions and thus no single response was coded with codes from more than one category.

For all of the coding, Cohen's Kappa [133], was used to measure inter-rater reliability (IRR). For this purpose, I and the previously mentioned collaborator both coded all of the responses to each question. IRR was then calculated and disagreements were discussed and resolved by the two coders. During the process, statements that were coded differently were discussed to determine differences in the coders' interpretations and code definitions were updated as needed. If necessary, the data-set was re-coded after discussion. The resulting Kappa values, all indicating strong to very strong agreement, for each SRL process can be seen in Table 11.2. I then coded the entire set of qualitative data, after which Chi-Square was used to check for significant differences between groups for each SRL construct.



(a) (b) (c)

Figure 11.3: Demographic breakdown of participants

	Avg (B)	StDev (B)	Avg (A)	StDev (A)	Significance
Process	3780.93	1176.65	4478.71	1203.42	$\boldsymbol{p = .03}$
Aggregate	3826.21	1215.87	3928.43	951.14	$p = .63$

Table 11.3: The average performance score before (B) and after (A) reflection for both groups, with standard deviation.

11.2.2 Results

11.2.2.1 Demographics

Demographic data collected included age, gender, and years of experience. This information can be seen in Figure 11.3.

11.2.2.2 Performance

The average scores for overall performance (based on the amount of gold earned with reductions for deaths) can be seen in Table 11.3. While both groups improved their performance, only the group that reflected on the process visualization experienced a significant improvement according to paired T-Tests ($t(13) = 2.52$, $p = .03$). The effect size, calculated using Cohen's d, was .67, indicating a moderate effect.

To better understand how performance changed, I analyzed each of the individual gameplay metrics that could impact the performance score, with the exception of lost towers, as no participant ever lost a tower. The average creep score (CS), which measures how many creeps and jungle camps a player killed (two of the four ways a player could earn gold within the confines of the task) for both groups for both attempts

	Avg (B)	**StDev (B)**	**Avg (A)**	**StDev (A)**	**Significance**
Process	60.36	15.71	65.79	14.05	$p = .12$
Aggregate	60.36	19.67	63.43	15.59	$p = .35$

Table 11.4: The average creep score before (B) and after (A) reflection for both groups, with standard deviation.

	Avg (B)	**StDev (B)**	**Avg (A)**	**StDev (A)**	**Significance**
Process	.86	1.29	.07	.27	$\boldsymbol{p = .01}$
Aggregate	.43	1.09	.5	.52	$p = .41$

Table 11.5: The average number of deaths before (B) and after (A) reflection for both groups, with standard deviation.

can be seen in Table 11.4. While both groups exhibited an improvement, neither change was significant (all p¿.05).

The average number of deaths, which would result in a deduction from the overall score, for both groups for both attempts can be seen in Table 11.5. While the aggregate visualization group experienced a non-significant (p¿.05) increase in the number of deaths, the process visualization group experienced a significant decrease in the number of deaths according to a Wilcoxon Signed Rank Test (Z=2.44, p=.01). The effect size, calculated via Z statistic divided by the square root of the sample size, is .65, indicating a large effect.

The average number of destroyed towers, which was one of the four ways that players could earn gold within the confines of the task, for both groups for both attempts can be seen in Table 11.6. While the process visualization group saw a slight increase and the aggregate visualization group saw a slight decrease, neither result was statistically significant (all p¿.05).

	Avg (B)	**StDev (B)**	**Avg (A)**	**StDev (A)**	**Significance**
Process	.07	.27	.36	.63	$p = .07$
Aggregate	.14	.36	.07	.27	p .34

Table 11.6: The average number of towers destroyed before (B) and after (A) reflection for both groups, with standard deviation.

	Avg (B)	StDev (B)	Avg (A)	StDev (A)	Significance
Process	1.79	1.37	2.29	1.54	$p = .4$
Aggregate	2.29	1.98	2	1.71	$p = .62$

Table 11.7: The average number of kills before (B) and after (A) reflection for both groups, with standard deviation.

	Avg	StDev
Process Group	48	29.09
Aggregate Group	45	25.76

Table 11.8: The average degree of behavioral change from before and after reflection for both groups, with standard deviation. The difference in average degree of behavioral change between the two groups was not significant ($p = .67$).

Finally, the average number of kills, which was one of the four ways that players could earn gold within the confines of the task, for both groups for both attempts can be seen in Table 11.7. While the process group saw a slight increase and the aggregate group saw a slight decrease, these changes were not statistically significant (all p¿.05).

11.2.2.3 Change

The average measure of change, calculated using DTW as discussed previously, can be seen in Table 11.8. The averages indicate that players in both groups changed their behavior between attempts to a similar degree, but the difference between groups was not significant (p¿.05) meaning that neither group experienced a significantly greater change of strategy than the other.

11.2.2.4 Self-Regulated Learning

The counts for how many times each "Evaluation" code was applied across the data set can be seen in table 11.9. Chi-Square indicated significant differences between the two groups ($\chi^2(4) = 15.36$, p = .004). The effect size, calculated via Cramer's V, was .74, indicating a large effect.

The counts for how many times each "Attribution" code was applied across

Code	Process Group	Aggregate Group
Performance of Other	0	1
Their Scores	1	10
Correct Method or Strategy	9	3
Improvement During Gameplay	3	0
They Don't Know	1	0

Table 11.9: Code counts for the evaluation process of SRL. These differences were significant at $p = .004$.

Code	Process Group	Aggregate Group
Specific Technique	5	6
General Technique	3	5
Confidence/Ability	1	0
Focus/Concentration	2	1
Practice	1	0
Rhythm	1	1
Distractions	1	0
Other	0	1

Table 11.10: Code counts for the attribution process of SRL. These differences were not significant at $p = .67$.

the data set can be seen in table 11.10. “Effort” and “They Don't Know” were never applied by either coder. Chi-Square indicated that these differences were not significant (p¿.05).

The counts for how many times each “Adaptation” code was applied across the data collected before the players' second attempt can be seen in Table 11.11. “Effort”, “Distractions”, “Other”, and “They Don't Know” were never applied by either coder and “Confidence/Ability” was omitted as discussed above. Chi-Square indicated that these differences were not significant (p¿.05).

The counts for how many times each “Goal Setting” code was applied across the data can be seen in Table 11.12. Chi-Square indicated that these differences were not significant (p¿.05).

The counts for how many times each “Strategy Use” code was applied across the data can be seen in Table 11.13. Chi-Square indicated that these differences were

Code	**Process Group**	**Aggregate Group**
Specific Technique	3	6
General Technique	8	5
Focus/Concentration	1	1
Practice	0	1
Rhythm	2	1

Table 11.11: Code counts for the adaptation process of SRL. These differences were not significant at $p = .55$.

Code	**Process Group**	**Aggregate Group**
Outcome Goals	8	8
Process Goals	5	6
Other	1	0

Table 11.12: Code counts for the goal setting process of SRL. These differences were not significant at $p = .58$.

not significant (p¿.05).

The counts for how many times each "Goal Realization" code was applied across the data can be seen in Table 11.14. Chi-Square indicated that these differences were not significant (p¿.05).

The counts for how many times each "Self-Monitoring" code was applied across the data can be seen in Table 11.15. Chi-Square indicated that these differences were not significant (p¿.05).

Code	**Process Group**	**Aggregate Group**
Specific Technique	8	5
Visualization Strategy	1	0
Concentration Strategy	2	5
Technique and Concentration	3	3
Practice/No Strategy	0	1

Table 11.13: Code counts for the strategy use process of SRL. These differences were not significant at $p = .41$.

Code	Process Group	Aggregate Group
Yes	9	7
Partially	3	5
No	2	2

Table 11.14: Counts for how many players in each group realized their goals. These differences were not significant at $p = .69$.

Code	Process Group	Aggregate Group
Score	5	4
Technique	6	3
Do Not Know	1	6
Other	2	1

Table 11.15: Code counts for the self-monitoring process of SRL. These differences were not significant at $p = .17$

11.2.2.5 Trends Apparent in Qualitative Data

Further qualitative analysis beyond what is discussed previously is beyond the scope of this work. However, some relevant trends did emerge that I discuss here, as they help with understanding the significant impacts seen in the quantitative data. First and foremost, players' reflection responses (across both groups) often focused on deaths, and how those deaths impacted their performance. For example: "I would say, just by looking at the 5 deaths, that I did not collect as much gold as possible because of the time down for those deaths" (Participant 5, Process), and "I died 4 times [which was] a negative impact on my ability to get more gold" (Participant 21, Aggregate). Reflection responses also indicate that, while reflecting, players were often trying to identify what caused their deaths. For example: "So if I'd been focusing on the tower aggro while pushing, even though there was no enemy around, I could have not died and pushed more" (Participant 1, Process), and "I think I focused more on that than on the actual farming itself, which caused me to overextend and get that death in the lane" (Participant 12, Aggregate). When discussing what they would do differently, players in both groups often discussed changes related to what they thought had caused their death. For example: "obviously, at one point I died by overstaying, so in terms of things I could adjust or change I would definitely be more cautious about when to back or to know my limits basically" (Participant 7, Process), and "I didn't die, but I came close a couple of times, so I should probably look for better trades" (Participant 18, Aggregate). Overall this highlights that many players focus on identifying mistakes, what caused them, and how to avoid them, with deaths being the most apparent and detrimental mistake available to players within the confines of this task.

Players did also discuss the other gameplay elements in their reflections. For example, creep score: "because my champ, Garen, has that spinning E, [last hitting creeps] is pretty easy once that levels up" (Participant 4, Process), and "I moved to the middle lane early, because I could get some uninterrupted [creep score], without having to worry about Zyra" (Participant 12, Aggregate). The number of kills, or failure to secure a kill, was also commonly discussed, for example: "getting myself to level faster would've given me better control over being able to actually succeed in getting at least one kill on the Zyra" (Participant 2, Aggregate), and "goal was to just execute more as far as better harass and just killing. Last time I almost got kills twice which kept me from snowballing" (Participant 22, Process). In many cases, these discussions of gameplay elements (including deaths) emerged in relation to goals set for the second execution of the task, i.e.: "Honestly, I just wanted to not die" (Participant 9, Process), and "Mentally, in my head, I knew that I'd gotten 86 cs last time and I was trying to get that same cs or something better" (Participant 18, Aggregate). This illustrates that while mistakes were at the forefront of players' minds, thinking about their goals and considering their progress towards those goals, often in numerical ways, was also a part of their reflection processes.

There were also some trends that only emerged or were more apparent in one group than in the other. Among the aggregate group, after the second attempt at the task, there was a readily apparent sentiment that players' reflections after the first attempt led them down the wrong path. For example: "I think I accomplished those goals of being more aggressive because I changed the play-style and went for a more offensive build, but I don't know if it was to better effect" (Participant 8, Aggregate), and "I think it was even worse...I had some ideas that I thought would help but I couldn't apply them" (Participant 23, Aggregate). A trend more apparent among players in the process group was an awareness of the timing of certain events, often involving criticism that important events had occurred later than they should have. For example: "I think my performance was...delayed...because my first kill wasn't until a little more than halfway through" (Participant 20, Process), and "the timing of when I backed for items could've been more ideal. Thinking about how much gold I had and I probably could've had an earlier item" (Participant 7, Process). This is not entirely surprising given the timeline-like format of the process visualization, but still a trend worth noting. Overall this suggests that there were some differences in how the presentation of data impacted reflection beyond what is presented in the quantitative data. In the next sections, I discuss the primary takeaways from all of these findings.

11.2.3 Discussion

11.2.3.1 Performance

As can be seen from the results, the group that reflected on the process visualization significantly improved their performance from before to after reflection while the group that reflected on the aggregate visualization did not experience a significant improvement. The apparent suggestion from these results is that reflecting on the process-visualization did, to some degree, benefit players and help them improve their gameplay. To better understand how it may have benefited them, I examined the changes in gameplay metrics from their first to second attempts presented in Tables 11.4 through 11.7. As stated earlier, there were four metrics that could impact performance in terms of gold earned: deaths, kills, destroyed towers, and creep score, where only deaths were a mistake that would detract from one's performance score (lost towers would have detracted as well but, as stated earlier, no one ever lost a tower). Players were informed of the subtraction for deaths upfront. Of these four elements, only the number of deaths saw a significant change, where the process group significantly reduced the number of times they died in their second attempt, after reflecting on their first attempt.

In the previous section, when discussing the qualitative trends, I demonstrated how players often focused on the number of times they died, what caused those deaths, and what they could do differently to not die again. These statements illustrate that players recognized death as a detrimental mistake and that much of their reflection process was focused on avoiding that mistake. Although players also discussed other elements of gameplay including creep score, suggesting that players were also thinking about these elements, only the number of deaths saw a significant change. Further, despite both groups being similarly focused on reducing the number of deaths and determining what caused their deaths, only the process group succeeded in significantly reducing the number of times they died. In fact, while it was not significant, the aggregate group slightly increased the average number of deaths.

This implies that reflecting on a process visualization may help a player better recognize and avoid mistakes, perhaps by allowing them to better identify the cause of those mistakes. The earlier work in this dissertation, presented in parts 1 and 2, highlighted that identifying one's mistakes is one of the hardest parts of reflecting on (and improving through said reflection) complex gameplay [349, 346], and thus, this finding may prove valuable for helping computational support tools better meet players' needs. While this theory cannot be confirmed by the work I present here, I present this as an opportunity to be explored further in future work.

11.2.3.2 Self-Reflection

I also saw a significant difference in how participants evaluated their gameplay during the reflection step. Specifically, those participants in the process group more often evaluated their performance based on their use of the correct method or strategy whereas those in the aggregate group more often evaluated based on score alone, as can be seen in Table 11.9. This implies that being able to see process information will prompt players to think more about their strategic decisions. While this may seem obvious, it is valuable to know that players are at least somewhat less likely to engage in this behavior if they do not see process information.

Cleary et al. [132], in their study of basketball players, defined an evaluation category they called "process" which contained evaluation based on technique or strategy and evaluation based on improvement over time, which was also slightly more common among our participants who reflected on the process visualization. Cleary et al.'s results showed that those who were better trained to self-regulate their learning, who, in turn, performed better, used evaluation techniques from the process category significantly more often. Based on the understanding that process-based evaluation results in higher performance, knowing that exposure to a process visualization may prompt it is a valuable finding for the design of future tools.

It should be noted, however, that there were no other significant differences in self-reflective processes, specifically in adaptation and attribution, which is unlike previous work [132, 343]. One interesting possibility is that, while the visualization of the data does not have a direct impact on how players reflect beyond evaluation, it may impact whether or not they are able to come to the right conclusions from that reflection. Specifically, attribution is about determining the cause of a mistake and adaptation is about deciding on a new strategy to pursue. In both cases, there is an element of accuracy, whether or not the right cause has been identified and if the right new strategy has been selected. For example, two people might attribute their separate mistakes to the fact that they got distracted during the task, but only one of them may be correct, and the other one may have actually failed due to their own skill level or another factor. As illustrated by this example, it is possible, that while both groups tended to attribute the same types of causes and adjust their gameplay in the same ways, only the process group was accurate in their judgments and that the aggregate group, ultimately, made the wrong choice. This theory is supported by the quotes from players we present in the previous section, where aggregate players suggested that their reflections led them down the wrong paths.

This suggests that players, regardless of visualization style, will think similarly about what they can do differently and how they can change their gameplay, a theory supported by the lack of significant differences in adaptation or the degree of change measured through DTW. However, as discussed earlier in this thesis, being aware of a mistake and being aware of what caused a mistake or how to overcome it are two very different challenges [349, 346]. Thus, it is possible that process-visualizations may have an impact on players' ability to come to correct conclusions from their reflective processes regarding the causes of their mistakes and how to overcome them, possibly by prompting process-oriented evaluation, even if how the attribute or adapt their mistakes does not change. In other words, a process visualization may help players come to more correct conclusions about the causes of their mistakes and how to avoid them in the future.

Future work is, however, needed to explore this further and confirm this theory, as the work in the previous part highlighted concerns regarding how players draw accurate conclusions from process-visualizations [350] and because evaluating the accuracy of players' attributions and adaptive strategies is beyond the scope of this work.

11.2.3.3 Other Elements of Self-Regulated Learning

Before moving on, it is worth noting that there was no significant difference in goal setting between the groups, as can be seen in Table 11.12. It seems somewhat surprising that, even after looking at a process visualization, the process group did not display a stronger tendency towards process goals. I acknowledge here that individual differences or a small sample size may be factors. Additionally, goal setting is not a self-reflection process but rather a part of forethought [756], thus it is possible that reflective visualizations would not directly impact it, which connects to an earlier discussion about how reflection phase computational support does not do a good job of connecting to the next forethought phase. However, it is additionally possible that players have a pre-disposition towards outcome goals due to community standards. Among esports communities, performance is often measured based on outcomes and milestones (number of kills, creep score, amount of gold, etc...). As members of these communities, our participants may have it ingrained within them to think about their gameplay in terms of these outcomes, leading to a tendency to set outcome goals, even when presented with process information. This may be a limitation that process-visualizations are not equipped to overcome.

I acknowledge, in relation to this, that players were prompted with the goal of earning as much gold as possible, which is ultimately an outcome goal, but I illustrated in the previous section that the goals players vocalized rarely directly mentioned gold. Previous work has suggested that process goals are correlated with skill and performance [132, 131, 343] and my earlier work highlighted how existing support for esports players is insufficient in supporting goal setting [345, 346]. This finding may support this claim, but future work is ultimately needed to explore this further.

Similarly, although not significant, there were some signs that there may be differences in self-monitoring after a player reflects on a process visualization. Self-monitoring, a performance phase process [756], is understudied in the existing literature on gaming, but my earlier work, discussed in part 1, found that players tended towards monitoring scoring metrics [345]. Here, I saw that those who reflected on the aggregate visualization had a tendency towards not self-monitoring or not knowing if they had, as can be seen from Table 11.15. As such, it is worth exploring further, in future work, whether or not process-visualizations during self-reflection can impact self-monitoring practices during future iterations of a task as this particular process is correlated with skill and performance [343].

11.3 Considerations for Using Process Visualizations to Improve Performance

Based on the discussion in the previous section, here I summarize the notable implications and key takeaways for future research and the development of process-visualizations for self-reflection.

11.3.1 Process Visualizations may prompt Process-Oriented Evaluation

The results suggest that exposure to a process visualization during self-reflection may prompt process-oriented evaluation techniques that focus on strategy usage or improvement over time. This is almost certainly because they give the players access to this information, whereas an aggregate visualization presents them with only score-oriented information. This highlights a potential benefit of process-visualizations as a tool for self-reflection, as process-oriented evaluation has demonstrated a strong correlation with high performance [132]. This process-oriented evaluation may have also been what led to participants' ability to identify the causes of their mistakes and the best ways to address them more accurately. While future work should explore this further, this makes a good argument for the inclusion of process-visualizations in retrospective visualization systems for esports.

11.3.2 Process Visualizations may lead to more Accurate Reflection

The results indicate that exposure to either a process or aggregate visualization will result in similar self-reflective processes, suggesting that, ultimately the presentation of data may not matter so much for the kinds of reflection that occur. However, the results also suggest that reflection alone is not enough and that the conclusions that the player comes to from the reflection must be accurate. Specifically, with regards to attributions and adaptations, simply attributing the cause of a failure may not matter as much in the grand scheme of things as whether or not that cause was correctly attributed. Similarly for adaptation, a new strategy may matter less than the correct new strategy. While future work is needed to confirm this theory, the findings suggest that, by reflecting on a process visualization, players were able to come to more accurate conclusions about the causes of their mistakes and how they could avoid them when they repeated the task, leading to an improvement in performance. This suggests that process-visualizations may, in fact, help players form more accurate causal relationships about their gameplay, addressing the shortcoming discussed in part 2 and supporting their future use.

11.3.3 Expert-like Input might be Necessary to reach Accurate Conclusions

Given the implication that reflection alone is not enough to improve performance, and that accurate conclusions matter, future retrospective visualization systems, whether they use process-visualizations or not, will need to explore ways to guide players to accurate conclusions. In the earlier parts of this dissertation, I illustrated how players often go to more experienced friends of teammates to seek input on how to improve [346], and thus, it may be possible to address this by incorporating intelligent input that helps players reflect on their performance the way a real expert would help them. As such, future research should develop a better understanding of how players seek expert input and ways to incorporate such co-regulated learning [273] into retrospective visualization systems. While I took a first step at exploring this topic (discussed in part 1 of this dissertation) there is a great deal of room for future work.

11.3.4 Process-Oriented Reflection may Influence Self-Monitoring Behavior

Previous work has highlighted that self-monitoring is a critical element of the performance phase of SRL [343, 132] that is often not performed unless prompted [333] and, as illustrated by the studies in the earlier parts of this thesis, not well understood or perceived by complex game players [345, 346]. Though not significant, the results did show a slight difference in self-monitoring between the two groups after reflection, where those who reflected on the process visualization demonstrated a higher tendency to engage in the behavior. While I can make no claims based on these results alone, this does suggest that future work should explore this topic further to determine whether or not there is a connection. This is especially worth exploring given that previous work has demonstrated a connection between self-monitoring and performance [132].

11.3.5 Process Visualizations do not necessarily prompt Process Goals

In part 1 of this thesis I discussed how goal setting was one of the only SRL processes that were engaged significantly differently across expertise levels in League of Legends [345] and in part 2 I illustrated how goal setting support is largely non-existent among existing computational support tools [346]. The results here, further suggest that reflecting on a process visualization will not necessarily prompt players to set process goals, which are associated with higher performance [132]. As such, goal setting, and supporting the setting of appropriate goals, which is necessary to prevent frustration [88], remains an open problem within the complex gaming domain and something that future work should focus on more closely.

Part VI

Conclusion

Chapter 12

Summary of Contributions

In this dissertation, I leveraged the theory of Self-Regulated Learning, which describes ways in which people can self-direct their own learning process in the absence of an educator, to understand how players learn to play games. Building upon this understanding, I explored ways we could design computational assistants to support Self-Regulated Learning in complex games, such that more players may succeed at high-level play, even in the absence of access to social learning opportunities. Specifically, by leveraging the Cyclical Phase Model, as its three phase understanding of learning maps well to gameplay cycles, in this thesis, I have explored, through user centric research, a number of topics related to how players learn, the state of the art of computational support for learning, and how we can improve it. In this conclusion, I summarize the contributions of each thrust of this work.

12.1 Thrust 1: Studies of Self-Regulated Learning in Complex Games

The goal of this thrust was to develop a stronger understanding of how players learn to play complex games. While there exist plenty of knowledge on the types of skills possessed by experienced players, how those skills are gained, and how that skill gain process can be supported, are under-studied. This thrust, thus, consisted of three studies aimed at answering the question "How do players engage Self-Regulated Learning skills in the context of learning and improving at play?" Two studies explored this in solo contexts and the third in a group context, aiming to understand the extent to which SRL is relevant to learning to play complex games and to highlight how important the role of others is in this process, to better motivate the development of tools that could fill that role given an absence of or lack of access to a community.

The results of these three studies provide the following contributions:

- An empirical understanding of the activities engaged by players attempting to learn and improve in the domain of esports and the challenges they face in that process. These are expected to generalize to other, similar types of complex games.
- A mapping of each activity and challenge to the SRL skills encompassed by the Cyclical Phase Model or to skills that are interrupted by the challenge at hand, indicating that players do leverage SRL when learning gameplay, even when not consciously aware of the theory.
- Implications for the design of computational assistants that could support learning and address challenges, informed by those results.
- An empirical, first exploration of social learning in the context of esports using the theory of Co-Regulated Learning as a lens that highlights key themes in how esports players interact with each others' learning processes. These are expected to generalize to other, similar types of complex games and gaming contexts.
- An empirical understanding of the significant role that others play in an individual's learning process, based on those results, motivating the need for work on systems that can fill that role when a community is neither available nor accessible.
- Concrete evidence of the differences in how self-regulated learning behaviors are engaged across different skill levels in League of Legends, which likely generalize to other esport games and complex games of similar design.

- Empirically illustrated differences in how self-regulated learning manifests across skill levels in esports when compared to traditional sports.
- A theory of the role of computational UI elements and support tools in teaching, supporting, and facilitating self-regulatory behaviors in the context of esports games.

12.2 Thrust 2: Supporting Self-Regulated Learning Skills in Complex Games through Computational Support

Informed by the results of the first thrust, this second thrust was built on the understanding that computational support tools played a key role in supporting and facilitating self-regulated learning behaviors in complex games. However, existing work had yet to develop any understanding of the state of the art of computational support for complex games or develop concrete design requirements from the perspective of SRL, necessary to know how said support could be improved and what direction research should move in next. Thus, the goal of this thrust was to develop a more informed understanding of how computational support tools supported self-regulated learning in Complex Games and how this could be improved by better understanding what it was that players wanted from these tools. This thrust consisted of two studies aimed at exploring the question “How do data-driven tools support self-regulated learning skills in complex games?”

The results of these studies provide the following contributions:

- A theoretical mapping of the Cyclical Phase Model to esports gameplay such that pre-game is understood to be the forethought phase, in-game is understood to be the performance phase, and post-game is understood to be the reflection phase.
- A taxonomy of nine interventions provided by existing computational assistants for esport games.

- A map of the state of the art of computational assistance for SRL (based on CPM) in esports that details how the nine interventions are leveraged by existing tools in terms of when, before, during, or after gameplay, they are offered to players.
- User-derived design requirements for computational assistants for esports.
- Explicit gaps in existing support and concrete opportunities for addressing them through future research and improved design considerations derived through the comparison of the design requirements to the map.

12.3 Thrust 3: Making Sense of Visualizations of Process

Informed by the results of the previous thrust, thrust 3 understood that players needed more causal information during self-reflection to better evaluate their performance and make accurate attributions of failure. Specifically, players articulated that it was typically very difficult to understand why they were not advancing of why they had experienced a certain undesirable outcome and that, often, this was where they would need to seek the aid of others in order to progress in their learning process. Thus I identified this as an opportunity to improve computational support for SRL in complex games. Based on the review of existing tools, I discovered that existing tools used primarily aggregate data visualization, which was helpful in that it could be easily understood, but did not provide players with the causal information they desired. Thus, I proposed the use of process visualizations as a way to improve SRL support in computational tools. Process visualizations are, however, much more complicated in terms of their visual appearance. Thus, in this thrust, I sought to develop an understanding of how players made meaning of such visualizations of their own and others' performance data in order to better inform the design of human-readable process visualizations that could be shown to players. Thus, this thrust consisted of two studies exploring the question: "How do players of complex games extract meaningful insights from visualizations of process?"

The results of these studies provide the following contributions:

- An interaction taxonomy for spatio-temporal gameplay data that highlights the information that players rely on when making sense of gameplay data.
- A process model for how the interactive activities in the taxonomy and information gleaned are leveraged by players extracting meaning from others' spatio-temporal gameplay data.
- An understanding of how the extraction of meaning from gameplay data, and especially process-oriented visualizations of gameplay data varies from traditional meaning-making in information visualization.
- A concrete approach to visualizing gameplay data as a process visualization based on abstractions of gameplay actions.
- An interaction taxonomy for process visualizations of gameplay data that presents distinct activities that players engage in when making sense of the visualization and highlights what information is most important to them.
- Two process models representing two sense-making methods for how the interactive activities in the taxonomy are leveraged by players extracting meaning from process visualizations of gameplay data.
- Considerations for when and how to design and implement process visualizations to support players in complex games.

12.4 Thrust 4: Reflecting on and Learning through Process Visualizations

The results of the previous thrust demonstrated the versatility of process visualizations, which generalize across numerous types of games regardless of genre or presence of a spatial component. Therefore, in this final thrust, I argued for their inclusion in computational support tools as a way to better support self-regulated learning processes during the self-reflection phase, discovered in earlier thrusts as a critical element of learning in complex games. However, earlier thrusts had not highlighted the benefits of process visualizations beyond theoretical claims. Thus, the goal of this thrust was to examine, empirically, whether or not, and in what ways, a process visualization impacted player reflection and performance. As such, armed with an understanding of how meaning is made from process visualizations, this final thrust consisted of two studies that explored the question "How do process visualizations of one's own and others' gameplay data impact self-reflection and learning?"

The results of these studies provide the following contributions:

- An empirical demonstration of the impact of reflecting on others' process information on willingness to adapt and quality of reflection when compared to only one's own.
- Design considerations for the inclusion and presentation of others' data in a process visualization based on these empirical results.
- An empirical demonstration of the impact of reflecting on process visualizations on self-reflection processes when compared to an aggregate visualization.
- An empirical demonstration of the impact of reflecting on process visualizations on performance when compared to an aggregate visualization.
- An understanding that process visualizations may better support the derivation of accurate conclusions from a reflective process, thus laying the foundation for their future study and use in commercial tools.
- Design considerations for the inclusion of process visualizations in computational support tools.

- Concrete opportunities to explore the utility of process visualizations for complex games further in future work.

Chapter 13

Limitations and Future Work

In this chapter I discuss the limitations of this work in terms of the theoretical limitations of Self-Regulated Learning and the Cyclical Phase Model as well as the methodological limitations of this work itself. I additionally discuss what these limitations mean for future work

13.1 Limitations of Self-Regulated Learning and the Cyclical Phase Model

13.1.1 The Scope of the Cyclical Phase Model

While SRL is a long-standing, well-tested, and empirically proven theory of learning that lends itself well to gaming contexts, here I acknowledge its limitations and how they relate to this work. In their 2011 meta-analysis of SRL literature, Sitzmann and Ely [624] highlighted a series of areas where SRL research had room to improve its understanding of learning. One such limitation they discuss is the focus on pre, mid, and post-training and the need to understand how self-regulation evolves and changes over time in a more fluid manner. I acknowledge here that this work, which relies on the Cyclical Phase Model and its understanding of the three phases, falls within the restraints of this limitation.

As discussed in a previous section, the three phases of CPM map explicitly to the well-established phases of gameplay: pre, mid, and post-game. However, these are likely not the only moments in which the cycle of planning, taking action, and evaluating the outcome can be experienced within a game. On a more granular level, before taking any action, a player devises a plan, and after the action is complete, reflects on the outcome. This information will inform the next decision they make, and this is likely to occur repeatedly over the course of gameplay, suggesting that the three cycles happen repeatedly in a nested manner within the larger cycle of pre, mid, and post-game. While defining and examining the nature of this phenomenon, and developing a concrete framework for Self-Regulated Learning is beyond the scope of this thesis, this is a topic I aim to explore further in future work.

Building on this discussion of the scope of the CPM cycle and the timing of behavior, CPM, as a model for SRL, does not account for the existence of reflection during the completion of a task, which would be considered the performance phase within the confines of the model. As a part of his 1984 book "The reflective practitioner: How professionals think in action" [584], Schon describes two types of reflection: in action and on action. While the work detailed in this dissertation focused on reflection on action (occurring after an event has completed), reflection in action (occurring during the event) is not properly accounted for by the Cyclical Phase Model itself. The model does allude to reflection occurring during performance with reference to behaviors such as monitoring progress towards goals and adjusting strategies [756], but does not explicitly discuss reflective behaviors (such as evaluation or adaptation) occurring during this phase.

Within the context of this work, study participants alluded to the existence of performance phase, in-the-moment, reflective practices, and thus, future work may benefit from further developing our understanding of the performance phase CPM and how the behaviors it encompasses occur within games, specifically focusing on those meta-cognitive behaviors related to reflection during the activity. It is, however, known that observing and measuring meta-cognitive behavior during the performance phase is a difficult task partially due to the risk of disrupting the activity and partially due to learners often not being aware of their meta-cognition during this phase unless explicitly prompted about it [333]. Thus there also remains a methodological challenge for how best to address these questions in future work.

13.1.2 Accounting for Individual Differences and Social Interaction

Another consideration in the SRL literature is its connection with other factors or traits and the extent to which the external traits impact the execution of SRL skills and, in turn, learning. For example, Wolters and Hussain [725] found a connection between grit and the execution of SRL skills and Reddy et al. [550] discuss how those with attention-deficit hyperactivity disorder (ADHD) struggle with self-regulation. This suggests that SRL, as a learning theory, is not applicable to every learner or, in the case of this work, every player and suggests that computational support tools that support players based on SRL theories or that are designed to support SRL may not be applicable to everyone. Even within the studies presented in this dissertation there were participants who did not express the same desires from an assistant, with some wanting explicit grades that they would be immediately willing to accept and use as a basis for adaptation and others distrusting any conclusion they did not come to themselves.

Further, cultural differences may also impact how players learn and the activities that they engage when doing so. While this applies to geographical cultural differences, considering that players around the world may not all learn or accept assistance in that learning in the same way, it also applies to cultural differences among player communities. Within this context, the theory of communities of practice becomes relevant. According to the theory, people who share a passion or interest will learn to engage with or perform it and learn to do so better over time through continued interaction with one another [378, 717]. This concept applies to this work not only in the context of the exploration of Co-Regulated Learning discussed in Part 1 but also to the ways in which participants discussed SRL or were observed to engage it in solo learning contexts throughout the dissertation.

As is illustrated in Part 1, primarily in Chapter 4, learning in complex games rarely occurs in a vacuum and players will often leverage the insights and expertise of others in order to learn, progress, and improve. In Chapter 4, this is explored empirically in the context of esports teams, which, by definition, become a community of practice as they are a group of individuals with a shared interest who develop through continued interaction with each other around that interest. Further, within this structure, members of the community, the esports team, often lower the learning curve for newcomers, as illustrated by the themes discussed in Chapter 4, meaning that esports teams fulfill one of the ways that the literature recognizes that communities of practice improve performance [390].

This connection links to the aforementioned limitation of SRL in terms of how it connects to individual differences. Learning may not only occur differently among individual players of complex games but among communities of practice as a whole, where players who belong to a team or club or guild and play together frequently may exhibit certain SRL skills but not others and this may be different from those skills exhibited by a different community. These differences may or may not coincide with culture, age, gender, or geographic location among other demographic characteristics. It is also likely that these communities extend into online spaces, where influencers, streamers, and forum writers form communities of practice around themselves and influence the learning of their followers, as online communities of practice are already heavily discussed in existing literature [632, 30, 266, 565]. The possibility that SRL and learning in general function differently in different communities, thus needing different support considerations, also poses an interesting juxtaposition to the tendency for the introduction of computational coaching tools to motivate players to play similarly to one another [326].

A full examination of individual differences, communities of practice, and how they influence once another and individual players' execution of SRL skills is, ultimately, beyond the scope of this thesis. However, in future work, I hope to explore differences in SRL in the context of complex games across different groups, with care taken to understand how these differences intersect with communities of practice both in person and online. Especially in the case of esports, which have achieved global appeal, understanding how learning occurs across different groups, and how individual differences may impact acceptance of a computational assistant, is critical to the long-term applicability of the results of this thesis.

13.2 Limitations of this Work

Here, I acknowledge several methodological limitations of this work. First, the work in thrusts 1 and 2 looked exclusively at esports games. This decision was motivated by a number of factors: an abundance of data within the domain, the presence of many existing computational assistants within the domain due to the previous point, the abundance of user experiences with computational assistants within the domain due to the previous two points, esports global popularity and increased use in serious and high-impact domains, and personal interest. While this means that the results derived in these first two thrusts illustrate SRL and how it intersects with computational support in the context of one of the most complicated complex gaming contexts, I acknowledge that there are questions of generalizability.

While CPM, as illustrated earlier, is applicable to games across genres, the details of how players engage with the learning processes it encompasses and how these are influenced by computational assistance may vary by genre. These questions of generalizability are especially of interest when looking at solo gaming contexts and the context of games that do not have large communities surrounding them (as is the case for many serious games) where players may view external support differently. Future work is necessary to ensure that the findings of these two thrusts generalize properly to these other types of complex games. Players' responses during the studies conducted in these two thrusts, however, suggested some degree of generalizability, which is promising for future work.

Second, I acknowledge that, in thrust 3, I only looked at two types of visualizations of process, one of which was carried forward into thrust 4. There certainly exist other ways to visualize gameplay data in a process oriented manner that are not accounted for within the scope of this dissertation work. Thus, here, I acknowledge that this work is limited in this capacity and that the approach chosen may not be the single most effective one. That being said, establishing that a process visualization of one type has an impact on self-reflection and performance in at least one type of complex game sets the foundation for future work to explore this further.

Additionally, given the similarities between the two taxonomies developed in the two studies in thrust 3, it seems promising that sense-making, and the impact that the insights extracted from that sense-making process, may generalize across visualization styles. This is promising as far as assuming, therefore, that the insights derived within this thesis regarding the impact of these visualizations on reflection and gameplay performance may generalize beyond just the visualizations seen within this thesis.

Connected to this consideration, I acknowledge that, in thrust 4, I leveraged a specific approach to the design and presentation of the process data. Previous work [274, 703] has found that aesthetics can have an impact on how data is viewed and interpreted by users and thus it is possible that the presentation of the information and the design of the visualization itself had an impact on how players reflected and, ultimately, on the results. While an exploration of all possible designs of process visualizations is beyond the scope of this work, future work can examine the extent to which aesthetics impact self-reflection, adaptation, and learning when reflecting on a process visualization.

Finally, I additionally acknowledge that the quantitative data collected in thrust 4 was collected from relatively small samples and that there are chances that the results presented here may change when the studies are repeated on larger sample sizes. I emphasize here that these sample sizes are consistent with previous work on SRL [343] as well as the ACM CHI standard [96]. Thus, in the context of this thesis, the sample sizes for both studies are enough to accept the results as valid. Nevertheless, future work is needed to confirm this and these results are thus presented as a stepping stone upon which this future work may build.

13.3 Additional Considerations for Future Work

Here I discuss additional considerations for future work that are not necessarily connected to a concrete limitation but are informed by trends that emerged throughout this work.

Over the course of this dissertation work, a number of considerations have emerged regarding the fair and ethical existence of computational assistance for complex games. In general, there are increasing concerns regarding the handling and sharing of player data [746, 599] that have yet to be properly addressed and resolved. During this thesis work, players often expressed concerns over the competitive fairness of the tools being studied and whether or not they would negatively impact their performance or expose them to toxicity within the community. Additionally, the concept of data literacy [106] raises questions about the equity of these tools and whether or not they could support all players in a fair manner or if a bare minimum amount of data literacy is required for the support to be effective. While exploring these questions of ethics was beyond the scope of this thesis work, this is something I hope to look into further in future work.

Additionally, as I have said multiple times throughout this thesis, games are an appropriate proxy for testing and deriving early requirements for UX of AI and computational support systems for similarly complex high-impact domains, such as disaster response. While exploring generalizability in this direction was beyond the scope of this thesis, in future work, I plan to build collaborations that allow me to take the findings of this work outside of games and examine how they may be adapted and applied in high-impact domains. Based on the findings of such work, my goal is to build better AI assistants for supporting learning and performance in complex tasks.

Finally, I acknowledge here that these tools are not meant for everyone and using them should not be a requirement. Many players enjoy the sensation of overcoming difficult challenges on their own and it varies from game to game and genre to genre. These tools are meant to be an option to players who feel they need them or simply want to use them, but games should not be designed to require their use as third party support. The desires of game designers should also be considered when developing these tools and their opinions as to whether or not they should be used with their games should be accounted for. As such, there are contexts in which other methods of supporting learning will be necessary and future work should explore these alternatives.

Chapter 14

Conclusion

In this chapter I will return to the problem I set out to address at the beginning of this dissertation and discuss how the work here addresses it and what new questions have emerged as a result of this work.

14.1 The Problem

Complex games have proven benefits for players and are seeing increased industry success as well as integration into serious, high-impact domains. In some cases, existing games developed for entertainment are being leveraged towards real world development and benefit, such as with esports being integrated into schools to help students learn teamwork, problem solving, and emotional regulation skills that transfer to academic contexts [727, 673, 479]. In other cases, gamified learning and training are seeing success as engaging ways to teach people new skills in both academic and work environments [321, 567].

The use of these games is, however, limited in that they are notoriously difficult to learn and master, making high-level complex gameplay inaccessible to many people. For many players, the means of overcoming this hurdle is to leverage the knowledge, aid, and guidance of others. However, communities do not exist for every game, especially for many serious games, and are not welcoming or accessible to every player, especially within domains such as esports. This results in situations where the benefits of complex gameplay are not available to all players, where games developed for serious purposes may be completely ineffective at their intended use, and diversity issues where those most capable of succeeding as high-skill-level players are those already welcomed by the community, if one exists.

14.2 Understanding Learning

As a first step to addressing this problem, I sought to understand how players in complex games learn to play. This was done initially through the two qualitative interview studies that, specifically, aimed at creating a formative theoretical understanding of learning in esports, a gaming context that encompasses some of the most complicated complex games. The outcomes of these two studies were a comprehensive taxonomy of the activities and challenges engaged and faced by esports players when trying to move up the ranks and a detailed set of themes surrounding social learning within the context and how players help each other learn and improve. The third study took a mixed-methods approach to understanding the use of SRL skill by players more specifically and, when combined with the results of the first study, provide an interesting insight into how SRL skills are prompted and potentially supported by esport games and their design elements.

From the first study we can see that SRL practices such as evaluating performance, setting goals, and executing plans are commonly leveraged by players when they are learning. Of interest here is that many of the players themselves did not explicitly use this terminology or suggest any awareness of the concept of SRL during the interviews. From the third study, we further see that players of across skill levels tend to leverage SRL skills to equivalent degrees. Combining these two sets of results, I propose that players who engage in esports play may be learning SRL skills simply through their interactions with the games. I further propose that the computational support systems within the games themselves, i.e. the scoreboards or post-play statistics, are further prompting and supporting the adoption and use of these skills.

Going beyond this work, for the future development and use of complex games, this suggests that SRL, as a learning theory, can be used as a foundation upon which learning and support systems may be structured. These do not necessarily have to be computational systems, but could be training regimens or written guides that focus on getting new players comfortable with the cycle of SRL that is apparently common in complex games. Further, it suggests that when developing complex games for serious and high-impact domains, the inclusion of computational systems with information similar to what is provided by those present in esports games may aid players in progressing through skill levels and thus enhance the success of the game. There are also implications for the gamification of learning technology systems that may be able to leverage design insights from esports systems, though further discussion of this point is beyond the scope of this work. Finally, these findings also suggest that complex games, especially esports, may be leveraged as a means for teaching SRL skills that could transfer to other domains.

The results of the third study from Thrust 1, however, also leave an open question not addressed by this work. It was found that more advanced players exhibited better SRL skill usage in the forethought phase and I hypothesized that this was due to this phase having no apparent support systems for SRL within the game itself. I then suggested that advanced players are likely learning the skills they use from teammates and coaches whereas novice players are less likely to have such contacts. Combined with the results of the second study, looking at social learning, this emphasizes the significant role of others on learning and how important it is for players to have access to some form of guide or input, motivating the need for computational assistants pursued in the rest of the thesis. This thesis did not, however, explore ways of developing better support for the forethought phase of SRL in complex games, instead focusing on the self-reflection phase. Thus, there is an opening that can be filled by future work in addressing how to best support and teach forethought phase SRL skills to complex game players and how these impact overall performance.

14.3 Understanding Self-Regulated Learning as a Social Activity

Before moving on to the next section, I follow the discussion of learning and understanding SRL in games with a bit more discussion of the social components of SRL, specifically the Cyclical Phase Model. While the model, as it is proposed and often discussed in previous work, does not account for social interaction, the results of this thesis suggest that, at least in the context of games, interaction with others plays a significant role. Specifically, participants throughout would explicitly or implicitly refer to coaches, teammates, friends, communities, and instructors who knew more and could aid them by identifying errors, suggesting changes, evaluating performance, and tracking improvement for the player. While the theory of Co-Regulated Learning was leveraged to better understand these phenomena, it remained clear that three phases of learning were fundamental to learning in, at least, esports games. These three phases, before, in, and post game which players understand as playing different roles in their learning process, even if they do not know what CPM is, are not discussed prominently within the existing work on CoRL. Thus, there are two theories relevant to learning in complex games that, at this time, do not intersect sufficiently.

Overall, this suggests that, in the future, for work in this area to advance, a new model of SRL may need to be proposed, one which combines the Cyclical Phase Model with Co-Regulated Learning to account for an captures how social interaction occurs across the three phases of learning. Such a model would certainly be of great benefit to the domain of complex games, where, as this thesis demonstrates, learning is understood to be both a social activity and one that maps across three phases in relation to when the learning task, the gameplay itself, occurs. In the process of generating such a novel understanding would come the need to further, and in more detail, examine the role of a mentor or instructor within game learning contexts. Such an individual may be a coach, teacher, or tutor and their actions may play a critical role in how learning is Co-Regulated. Understanding how this occurs better in gaming contexts and updating the models of SRL accordingly can lead to the development of, not only better tools, but better implementation of gamified learning.

It may be possible, additionally, that such an advanced model would be of benefit to learning sciences more broadly, as students in other context likely also rely on social interaction when adopting a new task. Social SRL is largely understudied overall, with only Hadwin et al.'s [497] work being prominently cited on the topic. In this regard, do I hope that the work presented in this thesis may advance our understanding of social learning within and beyond games.

One additional consideration of social learning that may need to be added to such a model of Social, Cyclical, Self-Regulated Learning is that of emotion. While the work in this thesis focused on the regulation of learning, how emotions and emotional regulation intersect with learning and especially social learning must be considered. To date, most theories and models of SRL feature little consideration for emotions, with Boekaerts' [75] "coping or well-being mode" being one of few acknowledgements of the role and potential setbacks of emotions in the learning process. This model, and most others, however, do not account for the social component, where both the learner and others may experience emotional states that can impact their learning. Especially in emotionally charged contexts, as complex games often are, expanding SRL theory to account for emotions, and especially social emotions, may further advance our understanding of the phenomenon.

14.4 Examining Computational Support

Following the earlier discussion, I argue that if we are to make complex gameplay, and its proven benefits, more accessible to more players and more effective in high-impact domains, we must better support players' ability to learn on their own. While there are many ways that we may support players in this capacity, the main take-aways discussed in the previous section suggest that computational assistants are a promising avenue that have already seen commercial success within some domains.

Computational assistants can, in a sense, act as an alternative to the aid of another, providing players with valuable input and guidance that can help them overcome obstacles in their learning journey.

Prior to this work, however, there was no comprehensive understanding of what computational assistants for complex games looked like, offered players, how they interfaced with SRL, or how well they aligned with what players desired from them. Thus was I motivated to examine these tools more closely in the hopes of discovering gaps in their support that may be filled to better address the needs and desires of players. Through a systematic review of existing tools I aimed to address the first issue, of not understanding how computational assistants aid players, and through data-driven retrospective interviews I aimed to address the second, of needing to know what players wanted or needed from these tools. The result is a taxonomy of interventions and concrete design requirements for the development of future tools.

Within the context of this dissertation, I identified the lack of causal information in existing tools as a means of improving computational support for SRL. Players felt that they often knew they had made a mistake, due to experiencing an unsatisfactory outcome, but were unsure what the mistake was. In the earlier work on learning, players articulated how one of the greatest advantages of having another person to talk to was having another set of eyes examine their performance and pinpoint what they could do better or what they explicitly did wrong. Here, players suggested that they wanted computational support to do the same, but unlike a real person, the tools, in their existing forms, were mostly only able to tell players that something was wrong, but not what. At the very least, players felt like the tools could do more to help them understand cause and effect. I identified this as an opportunity to advance the design of computational support tools for complex games to better meet players needs and better fulfill the role of supporting players in the absence of a community. Going forward, future tools developed both within and outside of complex games can aim to address this need for causal information as a way to set themselves apart from competitors and better aid players. In the context of serious domains, the inclusion of support systems that specifically focus on causal information may make the games used in these domains more effective.

Another prominent take-away from this work was the prominent lack of comprehensive support for the performance phase (i.e. during gameplay). Previous work has suggested that performance phase processes are often not engaged unless prompted [333] suggesting that this may be a detriment to players. As such, this is an opportunity for future work to explore performance phase support and develop more comprehensive systems for it. However, a prominent theme in the earlier work on social learning was that learning was meant to be performed before and after play but not during it and during play the player should be focused on winning. Combining this with the theme that players desired non-disruptive aid during the performance phase gleaned from the interviews around existing tools suggests that providing performance phase support is an intricate design challenge that must balance the need to support learning with players' desires to be able to focus on their in-game performance and chance of victory. Explicitly addressing this challenge is beyond the scope of this thesis, and as such I present this as an opportunity for future work that may further improve the design of computational support tools beyond what is presented in this work.

14.5 Advancing the Use of Process Visualizations

The culmination of the work discussed previously was the identification of process visualizations as an avenue to better support players learning process by presenting them with more causal information. This, of course, raised questions regarding how meaning is made from visualizations of process, which were explored in the third thrust of this work through think-aloud studies prompting players to extract insights from visualizations of gameplay. Following this, to understand the impact of process visualizations on players' self-reflection and performance, I conducted two experimental studies examining first adaptation and the impact of community data and then the impact of process more holistically when compared to an aggregate visualization. The outcomes of this work were two taxonomies of how meaning is made from spatio-temporal and process visualizations of gameplay data, respectively, and empirical results regarding the impact of process visualizations on SRL and performance.

More generally, from the taxonomies, I found that there were a number of differences in how players made meaning from game data compared to how other users made sense of visualized data in other domains. One of the most prominent differences was the importance of the context of the game from which the data was drawn, with players more reliant on understanding how that context influenced the visualized actions than users in other domains are reported to be. This reliance often resulted in assumptions about the reasoning of the player who took the action, and this generalized to process visualizations as well, where study participants often made assumptions regarding the skills and strategies of the player who took the visualized actions.

For the future use of process visualizations in complex games, this means that preserving context may be key to successful implementation but also, especially in the case of process-only visualizations, suggests limitations to their successful deployment. Specifically, familiarity with the game context may be critical to their success. In the study where players knew less about the game (Parallel) they were observed to be more likely to make relatively baseless assumptions and do less work to confirm whether or not they were true. This, of course, raises questions about the equitable use of these tools, discussed earlier, in that it suggests that they would only be of use to those players who are already very familiar with the games. As such, while I presented a set of design considerations for addressing these concerns here, beyond this dissertation their is an opportunity for future work to determine ways to effectively scaffold players in deriving accurate understandings of visualized data in process visualizations even if they are less familiar with the gaming context. This work may also generalize beyond games and apply to behavioral data visualization at large.

Beyond the taxonomies, the two experiments provide empirical evidence to support the proposed use of process visualizations to enhance learning by showing that they prompt adaptation through the inclusion of community data and improve performance through more accurate reflections. For the future design of computational support tools for complex games, these results suggest that the inclusion of process visualizations, which would arguably be more work to design, are a worthwhile endeavor. For future research, these findings lay the foundation for future work to examine more complicated scenarios, such as full multi-player gameplay, and to explore the impact of accurate reflection more fully. The study on adaptation, additionally, found that while the inclusion of community data did prompt adaptation, it only did so in one third of the group, suggesting the potential for a more effective method of getting learners to adopt the skill. These results are applicable to complex games but likely also generalize to learning technology at large that may similarly benefit from the causal information provided by a process visualization.

14.6 Wrapping Up

In summary, the goal of this work was to find a way to support players' learning processes in complex games such that they may better access the benefits of play, to make the games more effective in serious domains, and to hopefully address diversity issues brought about by the lack of accessibility to experienced others who may help someone learn. Through mixed-methods, user-centric research, I identified SRL, and specifically the Cyclical Phase Model, as a practice central to the learning process in games and found that the presence of other people had a notable impact on an individual's ability to learn and adopt SRL skills. In the absence of those people, computational support systems, including data visualizations, can fill the gap, but there is room for improvement, especially with regards to providing players with the causal information necessary to understand their mistakes and how to fix them. Process visualizations are one promising method to providing players with this causal information, which demonstrate promise as far as guiding accurate reflective and adaptive processes. Further, when combined with community data, they may further prompt adaptation in a complex gaming task.

Based on these findings, my overarching research question "How can we facilitate, enhance, and encourage self-regulated learning in the context of improving at complex game-play?" is answered largely via the statement "we can use computational assistants to encourage SRL in complex gameplay and can further facilitate and enhance the processes therein via the design and implementation of process visualizations within the tools." In doing so, I argue that complex games can be more accessible to more people and better leveraged in serious domains.

While one may wonder why visualizations are necessary for self-reflection when players can easily review recordings of their gameplay, I emphasize that reviewing recordings is a time consuming process that not everyone always has the time or patience for. A single visualization depicting action-by-action gameplay in a holistic manner can provide an at-a-glance overview of performance for reflective purposes that may occur quickly enough to fit comfortably in between adjacent games. As such the work herein is not meant to replace existing methods of review and reflection via video or simply memory, but to complement and enhance it.

Regarding the future design, development, and research of these tools, this dissertation provides user requirements, taxonomies, theoretical insights, design recommendations, and empirical results to inform and guide. Potential directions for future work include the examination of the forethought and performance phases and ways to enhance SRL support and computational assistant design for these phases, neither of which were the focus of this work, opportunities to explore social learning and communities of practice within complex gaming further in terms of how learning activities vary, and the need to understand the fairness, ethics, and equity of computationally assistive tools for gaming in terms of intersecting concepts such as data literacy, privacy, and accessibility. The empirical results presented in this thesis further lay the foundation for additional research into reflection in gaming and its impact on performance.

Bibliography

[1] Onni Aarne, Petrus Peltola, Juho Leinonen, and Arto Hellas. A Study of Pair Programming Enjoyment and Attendance using Study Motivation and Strategy Metrics. In *Proceedings of the 49th ACM Technical Symposium on Computer Science Education - SIGCSE '18*, volume 2018-Janua, pages 759–764, New York, New York, USA, 2018. ACM Press.

[2] Onni Aarne, Petrus Peltola, Juho Leinonen, and Arto Hellas. A study of pair programming enjoyment and attendance using study motivation and strategy metrics. In *Proceedings of the 49th ACM Technical Symposium on Computer Science Education*, SIGCSE '18, page 759–764, New York, NY, USA, 2018. ACM.

[3] Maurice Abi-Raad. Code of ethics? which one? *ACM SIGCSE Bulletin*, 31(2):73, Jun 1999.

[4] Simon Abramov, Alexander Korotin, Andrey Somov, Evgeny Burnaev, Anton Stepanov, Dmitry Nikolaev, and Maria A Titova. Analysis of video game players' emotions and team performance: An esports tournament case study. *IEEE Journal of Biomedical and Health Informatics*, 26(8):3597–3606, 2021.

[5] Abir Abyaa, Mohammed Khalidi Idrissi, and Samir Bennani. Learner modelling: systematic review of the literature from the last 5 years. *Educational Technology Research and Development*, 67(5):1105–1143, 2019.

[6] ACM Special Interest Group on Design Automation., IEEE Council on Electronic Design Automation., and IEEE Circuits and Systems Society. *Proceedings of the*

45th Design Automation Conference : [Anaheim Convention Center, Anaheim, CA, June 9-13, 2008 ; sponsored by: ACM/SIGDA ... [et al.]. IEEE, 2008.

[7] Donald Acton, Allison Elliott Tew, Kurt Eiselt, ACM. Special Interest Group on Computer Science Education, ACM, and ACM Digital Library. *Proceedings of the seventeenth Western Canadian Conference on Computing Education : Vancouver, British Columbia, Canada : May 4-5, 2012.*

[8] Liz. Adams, ACM Digital Library., and ACM. Special Interest Group on Computer Science Education. *Proceedings of the 16th annual conference reports on Innovation and technology in computer science education - working group reports.* ACM, 2011.

[9] Chiranjivi Adhikari, Avash Puri, Dipty Thapa, Rebishna Thapa, Sonam Magar, and GC Sunil. Application of social cognitive theory in obesity prevention: a rapid review. *Journal of Health and Allied Sciences*, 7(1):53–62, 2018.

[10] Ana Paula Afonso, Maria Beatriz Carmo, and Rafael Afonso. Visualeague: Visual analytics of multiple games. In *2021 25th International Conference Information Visualisation (IV)*, pages 54–62. IEEE, 2021.

[11] Ana Paula Afonso, Maria Beatriz Carmo, Tiago Gonçalves, and Pedro Vieira. Visualeague: Player performance analysis using spatial-temporal data. *Multimedia Tools and Applications*, 78(23):33069–33090, 2019.

[12] Khuloud Ahmad and Paul Gestwicki. Studio-based learning and app inventor for android in an introductory cs course for non-majors. In *Proceeding of the 44th ACM Technical Symposium on Computer Science Education*, SIGCSE '13, page 287–292, New York, NY, USA, 2013. ACM.

[13] Khuloud Ahmad and Paul Gestwicki. Studio-based learning and app inventor for android in an introductory CS course for non-majors. In *Proceeding of the 44th*

ACM technical symposium on Computer science education - SIGCSE '13, page 287, New York, New York, USA, 2013. ACM Press.

[14] Sabbir Ahmad, Andy Bryant, Erica Kleinman, Zhaoqing Teng, Truong-Huy D Nguyen, and Magy Seif El-Nasr. Modeling individual and team behavior through spatio-temporal analysis. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, pages 601–612, 2019.

[15] Joseph Ahn, William Collis, and Seth Jenny. The one billion dollar myth: Methods for sizing the massively undervalued esports revenue landscape. *International Journal of Esports*, 1(1), 2020.

[16] John O. Aikin. Computer literacy. In Intergovernmental Panel on Climate Change, editor, *Proceedings of the ninth SIGCSE technical symposium on Computer science education - SIGCSE '78*, volume 44, pages 8–12, New York, New York, USA, Nov 1978. ACM Press.

[17] Arjun R Akula, Changsong Liu, Sinisa Todorovic, Joyce Y Chai, and Song-Chun Zhu. Explainable ai as collaborative task solving. In *CVPR Workshops*, pages 91–94, 2019.

[18] Satu Alaoutinen. Effects of learning style and student background on self-assessment and course performance. In *Proceedings of the 10th Koli Calling International Conference on Computing Education Research, Koli Calling'10*, pages 5–12, 2010.

[19] Satu Alaoutinen. Evaluating the effect of learning style and student background on self-assessment accuracy. *Computer Science Education*, 22(2):175–198, 2012.

[20] Meghan Allen, Jessica Q. Dawson, Celina Berg, and Neil Leveridge. Insights from the application of universal design principles to support English language learners. *Proceedings of the 23rd Western Canadian Conference on Computing*

Education, WCCCE 2018 - Held in cooperation with the ACM Special Interest Group on Computer Science Education (SIGCSE), 2018.

[21] Cristina Alonso-Fernandez, Antonio Calvo, Manuel Freire, Ivan Martinez-Ortiz, and Baltasar Fernandez-Manjon. Systematizing game learning analytics for serious games. In *2017 IEEE global engineering education conference (EDUCON)*, pages 1111–1118. IEEE, 2017.

[22] Yasmeen Alufaisan, Laura R Marusich, Jonathan Z Bakdash, Yan Zhou, and Murat Kantarcioglu. Does explainable artificial intelligence improve human decision-making? *arXiv preprint arXiv:2006.11194*, 2020.

[23] Khaila Amazan-Hall, Jen Jen Chen, Kathy Chiang, Amanda LL Cullen, Mark Deppe, Edgar Dormitorio, Doug Haynes, Jessica Kernan, Kirsten Quanbeck, Morgan Romine, et al. Diversity and inclusion in esports programs in higher education: Leading by example at uci. *International Journal of Gaming and Computer-Mediated Simulations (IJGCMS)*, 10(2):71–80, 2018.

[24] Saleema Amershi, Maya Cakmak, William Bradley Knox, and Todd Kulesza. Power to the people: The role of humans in interactive machine learning. *Ai Magazine*, 35(4):105–120, 2014.

[25] Saleema Amershi, Dan Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi Iqbal, Paul N Bennett, Kori Inkpen, et al. Guidelines for human-ai interaction. In *Proceedings of the 2019 chi conference on human factors in computing systems*, pages 1–13, 2019.

[26] Erik Andersen, Yun-En Liu, Ethan Apter, François Boucher-Genesse, and Zoran Popović. Gameplay analysis through state projection. In *Proceedings of the fifth international conference on the foundations of digital games*, pages 1–8, 2010.

[27] Erik Andersen, Eleanor O'rourke, Yun-En Liu, Rich Snider, Jeff Lowdermilk, David Truong, Seth Cooper, and Zoran Popovic. The impact of tutorials on

games of varying complexity. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 59–68, 2012.

[28] Thomas A. Angelo and K. Patricia Cross. *Classroom Assessment Techniques: A Handbook for Faculty.* Office of Educational Research and Improvement, NCRIPT, Ann Arbor, Michigan, Sep 1988.

[29] Thomas A. Angelo and K. Patricia Cross. *Classroom Assessment Techniques: A Handbook for College Teachers.* Jossey-Bass, San Francisco, California, 1993.

[30] Jo Angouri. Online communities and communities of practice. In *The Routledge handbook of language and digital communication*, pages 323–338. Routledge, 2015.

[31] Kai Arakawa, Qiang Hao, Tyler Greer, Lu Ding, Christopher D. Hundhausen, and Abigayle Peterson. In situ identification of student self-regulated learning struggles in programming assignments. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education*, SIGCSE '21, page 467–473, New York, NY, USA, 2021. Association for Computing Machinery.

[32] Sarah L Ash and Patti H Clayton. The articulated learning: An approach to guided reflection and assessment. *Innovative Higher Education*, 29(2):137–154, 2004.

[33] Tapio Auvinen, Juha Paavola, and Juha Hartikainen. STOPS. In *Proceedings of the 14th Koli Calling International Conference on Computing Education Research - Koli Calling '14*, volume 2014-Novem, pages 25–34, New York, New York, USA, 2014. ACM Press.

[34] Reyyan. Ayfer, ACM Digital Library., and ACM. Special Interest Group on Computer Science Education. *Proceedings of the fifteenth annual conference on Innovation and technology in computer science education.* ACM, 2010.

[35] Eshwar Bachu and Margaret Bernard. A computer supported collaborative learning (cscl) model for educational multiplayer games. In *11th International*

Conference on e-Learning, e-Business, Enterprise Information Systems, and e-Government, Las Vegas, NV, 2012.

[36] Steve Olusegun Bada and Steve Olusegun. Constructivism learning theory: A paradigm for teaching and learning. *Journal of Research & Method in Education*, 5(6):66–70, 2015.

[37] Carole A. Bagley and C. Candace Chou. Collaboration and the importance for novices in learning java computer programming. In *Proceedings of the 12th Annual SIGCSE Conference on Innovation and Technology in Computer Science Education*, ITiCSE '07, page 211–215, New York, NY, USA, 2007. Association for Computing Machinery.

[38] Joseph Baker, Jeane Cote, and Bruce Abernethy. Sport-specific practice and the development of expert decision-making in team ball sports. *Journal of Applied Sport Psychology*, 15(1):12–25, 2003.

[39] Albert Bandura. The explanatory and predictive scope of self-efficacy theory. *Journal of Social and Clinical Psychology*, 4(3):359–373, 1986.

[40] Albert Bandura. *Social Foundations of Thought and Action: A Social Cognitive Theory*. Prentice-Hall, 1986.

[41] Albert Bandura. *Self-efficacy: The Exercise of Control*. W H Freeman/Times Books/ Henry Holt & Co., 1997.

[42] Albert Bandura. The evolution of social cognitive theory. *Great minds in management*, pages 9–35, 2005.

[43] Albert Bandura and Richard H. Walters. *Social Learning Theory*, volume 1. Prentice-Hall, 1977.

[44] Maria Bannert. Promoting self-regulated learning through prompts. *Zeitschrift für Pädagogische Psychologie*, 23(2):139–145, 2009.

[45] Maria Bannert and Peter Reimann. Supporting self-regulated hypermedia learning through prompts. *Instructional Science*, 40(1):193–211, 2012.

[46] Gagan Bansal, Besmira Nushi, Ece Kamar, Walter S Lasecki, Daniel S Weld, and Eric Horvitz. Beyond accuracy: The role of mental models in human-ai team performance. In *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, volume 7, pages 2–11, 2019.

[47] Erik Barendsen, Linda Mannila, Barbara Demo, Nataša Grgurina, Cruz Izu, Claudio Mirolo, Sue Sentance, Amber Settle, and Gabrielė Stupurienė. Concepts in K-9 Computer Science Education. In *Proceedings of the 2015 ITiCSE on Working Group Reports - ITICSE-WGR '15*, pages 85–116, New York, New York, USA, 2015. ACM Press.

[48] Lucy Barnard-Brak, Valerie Osland Paton, and William Y Lan. Profiles in self-regulated learning in the online learning environment. *International Review of Research in Open and Distributed Learning*, 11(1):61–80, 2010.

[49] Charles Daniel Batson. *Altruism in humans.* Oxford University Press, USA, 2011.

[50] Lina Battestilli, Apeksha Awasthi, and Yingjun Cao. Two-Stage Programming Projects. In *Proceedings of the 49th ACM Technical Symposium on Computer Science Education - SIGCSE '18*, volume 2018-Janua, pages 479–484, New York, New York, USA, 2018. ACM Press.

[51] Brett A. Becker, Paul Denny, Raymond Pettit, Durell Bouchard, Dennis J. Bouvier, Brian Harrington, Amir Kamil, Amey Karkare, Chris McDonald, Peter-Michael Osera, Janice L. Pearce, and James Prather. Compiler error messages considered unhelpful: The landscape of text-based programming error message research. In *Proceedings of the Working Group Reports on Innovation and Technology in Computer Science Education*, ITiCSE-WGR '19, page 177–210, New York, NY, USA, 2019. Association for Computing Machinery.

[52] Brett A. Becker and Keith Quille. 50 years of cs1 at sigcse: A review of the evolution of introductory programming education research. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education*, SIGCSE '19, page 338–344, New York, NY, USA, 2019. ACM.

[53] Marc Berges, Andreas Mühling, and Peter Hubwieser. The gap between knowledge and ability. In *Proceedings of the 12th Koli Calling International Conference on Computing Education Research - Koli Calling '12*, pages 126–134, New York, New York, USA, 2012. ACM Press.

[54] Susan Bergin and R. Reilly. The influence of motivation and comfort-level on learning to program. In *PPIG*, pages 293–304. Psychology of Programming Interest Group, 2005.

[55] Susan Bergin, Ronan Reilly, and Desmond Traynor. Examining the role of self-regulated learning on introductory programming performance. In *Proceedings of the First International Workshop on Computing Education Research*, ICER '05, page 81–86, New York, NY, USA, 2005. ACM.

[56] Nadine Bergner, Jan Holz, and Ulrik Schroeder. InfoSphere: An Extracurricular Learning Environment for Computer Science. In *Proceedings of the 7th Workshop in Primary and Secondary Computing Education on - WiPSCE '12*, page 22, New York, New York, USA, 2012. ACM Press.

[57] Margaret Bernard and Eshwar Bachu. Enhancing the metacognitive skill of novice programmers through collaborative learning. In *Metacognition: Fundaments, applications, and trends*, pages 277–298. Springer, Cham, CH, 2015.

[58] Gowri Betrabet Gulwadi. Using reflective journals in a sustainable design studio. *International Journal of Sustainability in Higher Education*, 10(1):43–53, 2009.

[59] Anastasia Lynn Betts and Meagan K Rothschild. Massively multiplayer online games as spaces for metacognition and self-regulated learning. In *Global Perspec-*

tives on Gameful and Playful Teaching and Learning, pages 78–104. IGI Global, 2020.

[60] Jyoti Bhardwaj. New Entrants' Expectations of the First Year Computer Science Experience in the Context of a New National High School Curriculum. In *Proceedings of the 2019 ACM Conference on Innovation and Technology in Computer Science Education - ITiCSE '19*, pages 85–91, New York, New York, USA, 2019. ACM Press.

[61] Katerine Bielaczyc, Peter L. Pirolli, and Ann L. Brown. Training in self-explanation and self-regulation strategies: Investigating the effects of knowledge acquisition activities on problem solving. *Cognition and Instruction*, 13(2):221–252, 1995.

[62] Reuben Binns, Max Van Kleek, Michael Veale, Ulrik Lyngs, Jun Zhao, and Nigel Shadbolt. 'it's reducing a human being to a percentage': Perceptions of justice in algorithmic decisions. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, page 377. ACM, 2018.

[63] Aaron E. Black and Edward L. Deci. The effects of instructors' autonomy support and students' autonomous motivation on learning organic chemistry: A self-determination theory perspective. *Science Education*, 84(6):740–756, 2000.

[64] A. F. Blackwell. First steps in programming: a rationale for attention investment models. In *Proceedings IEEE 2002 Symposia on Human Centric Computing Languages and Environments*, pages 2–10, Sep 2002.

[65] Alan F Blackwell, Marian Petre, and Luke Church. Fifty years of the psychology of programming. *International Journal of Human-Computer Studies*, 131:52–63, 2019.

[66] Suzan JE Blijlevens, Marije T Elferink-Gemser, Paul Wylleman, Kayan Bool, and Chris Visscher. Psychological characteristics and skills of top-level dutch

gymnasts in the initiation, development and mastery stages of the athletic career. *Psychology of Sport and Exercise*, 38:202–210, 2018.

[67] Joan Bliss, Mike Askew, and Sheila Macrae. Effective teaching and learning: Scaffolding revisited. *Oxford review of Education*, 22(1):37–61, 1996.

[68] Blitz.gg. Blitz App. https://blitz.gg/, 2021. Online; accessed April 15th 2021.

[69] Benjamin Samuel Bloom. *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. Longman, 2001.

[70] Marc Bloom. Self-regulated learning: Goal setting and self-monitoring. *The Language Teacher*, 37(4):46–51, 2013.

[71] Fran C Blumberg, Sheryl F Rosenthal, and John D Randall. Impasse-driven learning in the context of video games. *Computers in Human Behavior*, 24(4):1530–1541, 2008.

[72] Konrad Bochennek, Boris Wittekindt, Stefanie-Yvonne Zimmermann, and Thomas Klingebiel. More than mere games: a review of card and board games for medical education. *Medical teacher*, 29(9-10):941–948, 2007.

[73] Robert Bodily, Judy Kay, Vincent Aleven, Ioana Jivet, Dan Davis, Franceska Xhakaj, and Katrien Verbert. Open learner models and learning analytics dashboards: a systematic review. In *Proceedings of the 8th international conference on learning analytics and knowledge*, pages 41–50, 2018.

[74] Robert Bodily and Katrien Verbert. Review of research on student-facing learning analytics dashboards and educational recommender systems. *IEEE Transactions on Learning Technologies*, 10(4):405–418, 2017.

[75] Monique Boekaerts and Lyn Corno. Self-regulation in the classroom: A perspective on assessment and intervention. *Applied psychology*, 54(2):199–231, 2005.

[76] Monique Boekaerts and Lyn Corno. Self-regulation in the classroom: A perspective on assessment and intervention. *Applied Psychology*, 54(2):199–231, 2005.

[77] Niall Bolger and John Eckenrode. Social relationships, personality, and anxiety during a major stressful event. *Journal of Personality and Social psychology*, 61(3):440, 1991.

[78] Nigel Bosch, Yingbin Zhang, Luc Paquette, Ryan Baker, Jaclyn Ocumpaugh, and Gautam Biswas. Students' verbalized metacognition during computerized learning. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–12, New York, NY, USA, 2021. Association for Computing Machinery.

[79] Matt Bower. Virtual classroom pedagogy. *ACM SIGCSE Bulletin*, 38(1):148, 2006.

[80] Brian Bowman, Niklas Elmqvist, and TJ Jankun-Kelly. Toward visualization for games: Theory, design space, and patterns. *IEEE transactions on visualization and computer graphics*, 18(11):1956–1968, 2012.

[81] J. Bransford, A.L. Brown, and R.R. Cocking. *How People Learn: Brain, Mind, Experience, and School: Expanded Edition.* The National Academies Press, Washington, DC, 2000.

[82] John Bransford, John D Bransford, Ann L Brown, and Rodney R Cocking. *How people learn: Brain, mind, experience, and school.* National Academies Press, 1999.

[83] Virginia Braun and Victoria Clarke. Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2):77–101, 2006.

[84] Ann L. Brown. The development of memory: Knowing, knowing about knowing, and knowing how to know. volume 10 of *Advances in Child Development and Behavior*, pages 103 – 152. JAI, 1975.

[85] Peter Brusilovsky, Girish Chavan, and Rosta Farzan. Social adaptive navigation support for open corpus electronic textbooks. In *International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems*, pages 24–33. Springer, 2004.

[86] Peter Brusilovsky, I-Han Hsiao, and Yetunde Folajimi. Quizmap: open social student modeling and adaptive navigation support with treemaps. In *Towards Ubiquitous Learning: 6th European Conference of Technology Enhanced Learning, EC-TEL 2011, Palermo, Italy, September 20-23, 2011. Proceedings 6*, pages 71–82. Springer, 2011.

[87] Peter Brusilovsky and Eva Millán. User models for adaptive hypermedia and adaptive educational systems. In *The adaptive web*, pages 3–53. Springer, 2007.

[88] Robert C Brusso, Karin A Orvis, Kristina N Bauer, and Amanuel G Tekleab. Interaction among self-efficacy, goal orientation, and unrealistic goal-setting on videogame-based training performance. *Military Psychology*, 24(1):1–18, 2012.

[89] Kevin Buffardi, Pedro Valdivia, and Destiny Rogers. Measuring Unit Test Accuracy. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education - SIGCSE '19*, pages 578–584, New York, New York, USA, Feb 2019. ACM Press.

[90] Susan Bull. There are open learner models about! *IEEE Transactions on Learning Technologies*, 13(2):425–448, 2020.

[91] Barry Burd, Lecia Barker, Félix Armando, Fermín Pérez, Ingrid Russell, Bill Siever, Liviana Tudor, Michael McCarthy, and Ian Pollock. The Internet of Things in undergraduate computer and information science education: Exploring curricula and pedagogy. *Annual Conference on Innovation and Technology in Computer Science Education, ITiCSE*, pages 200–216, 2018.

[92] Quinn Burke and Cinamon Sunrise Bailey. Camp or College? In *Proceedings of*

the 50th ACM Technical Symposium on Computer Science Education - SIGCSE '19, pages 345–350, New York, New York, USA, 2019. ACM Press.

[93] Emily A Butler and Ashley K Randall. Emotional coregulation in close relationships. *Emotion Review*, 5(2):202–210, 2013.

[94] Zack Butler, Ivona Bezakova, and Kimberly Fluet. Pencil puzzles for introductory computer science: An experience- and gender-neutral context. In *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education*, SIGCSE '17, page 93–98, New York, NY, USA, 2017. Association for Computing Machinery.

[95] Zack Butler, Ivona Bezáková, and Kimberly Fluet. Analyzing rich qualitative data to study pencil-puzzle-based assignments in CS1 and CS2. In *Proceedings of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education - ITiCSE 2018*, pages 212–217, New York, New York, USA, 2018. ACM Press.

[96] Kelly Caine. Local standards for sample size at chi. In *Proceedings of the 2016 CHI conference on human factors in computing systems*, pages 981–992, 2016.

[97] Jennifer Campbell, Diane Horton, and Michelle Craig. Factors for success in online cs1. In *Proceedings of the 2016 ACM Conference on Innovation and Technology in Computer Science Education*, ITiCSE '16, page 320–325, New York, NY, USA, 2016. Association for Computing Machinery.

[98] Jennifer Campbell, Andrew Petersen, and Jacqueline Smith. Self-paced Mastery Learning CS1. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education - SIGCSE '19*, pages 955–961, New York, New York, USA, 2019. ACM Press.

[99] John L Campbell, Charles Quincy, Jordan Osserman, and Ove K Pedersen. Cod-

ing in-depth semistructured interviews: Problems of unitization and intercoder reliability and agreement. *Sociological Methods & Research*, 42(3):294–320, 2013.

[100] Mickael Campo, Diane M Mackie, and Xavier Sanchez. Emotions in group sports: A narrative review from a social identity perspective. *Frontiers in psychology*, 10:666, 2019.

[101] Mickaël Campo, Xavier Sanchez, Claude Ferrand, Elisabeth Rosnet, Andrew Friesen, and Andrew M. Lane. Interpersonal emotion regulation in team sport: Mechanisms and reasons to regulate teammates' emotions examined. *International Journal of Sport and Exercise Psychology*, 15(4):379–394, Aug 2017.

[102] Yingjun Cao, Leo Porter, Soohyun Nam Liao, and Rick Ord. Paper or Online? In *Proceedings of the 2019 ACM Conference on Innovation and Technology in Computer Science Education - ITiCSE '19*, pages 99–104, New York, New York, USA, 2019. ACM Press.

[103] Angela Carbone, John Hurst, Ian Mitchell, and Dick Gunstone. Principles for designing programming exercises to minimise poor learning behaviours in students. In *Proceedings of the Australasian conference on Computing education - ACSE '00*, pages 26–33, New York, New York, USA, 2000. ACM Press.

[104] Angela Carbone and Judy Sheard. A studio-based teaching and learning model in IT. In *Proceedings of the 7th annual conference on Innovation and technology in computer science education - ITiCSE'02*, page 213, New York, New York, USA, 2002. ACM Press.

[105] Andy Carle, Michael Clancy, and John Canny. Working with pedagogical patterns in PACT. In *Proceedinds of the 38th SIGCSE technical symposium on Computer science education - SIGCSE '07*, page 238, New York, New York, USA, 2007. ACM Press.

[106] Jake Carlson and Lisa Johnston. *Data information literacy: Librarians, data, and the education of a new generation of researchers*. Purdue University Press, 2015.

[107] Albert V Carron, Steven R Bray, and Mark A Eys. Team cohesion and team success in sport. *Journal of sports sciences*, 20(2):119–126, 2002.

[108] Adam S. Carter and Christopher D. Hundhausen. With a little help from my friends: An empirical study of the interplay of students' social activities, programming activities, and course success. In *ICER 2016 - Proceedings of the 2016 ACM Conference on International Computing Education Research*, pages 201–209. ACM, Inc, Aug 2016.

[109] Janet Carter, Jian Shi, Su White, Dennis Bouvier, Rachel Cardell-Oliver, Margaret Hamilton, Stanislav Kurkovsky, Stefanie Markham, O. William McClung, Roger McDermott, and Charles Riedesel. Motivating all our students? In *Proceedings of the 16th annual conference reports on Innovation and technology in computer science education - working group reports - ITiCSE-WGR '11*, page 1, New York, New York, USA, 2011. ACM Press.

[110] Milly Casey-Campbell and Martin L Martens. Sticking it all together: A critical assessment of the group cohesion–performance literature. *International Journal of Management Reviews*, 11(2):223–246, 2009.

[111] Michael E. Caspersen and Jens Bennedsen. Instructional design of a programming course. In *Proceedings of the third international workshop on Computing education research - ICER '07*, page 111, New York, New York, USA, 2007. ACM Press.

[112] Sven Charleer, Kathrin Gerling, Francisco Gutiérrez, Hans Cauwenbergh, Bram Luycx, and Katrien Verbert. Real-time dashboards to support esports spectating. In *Proceedings of the 2018 Annual Symposium on Computer-Human Interaction in Play*, pages 59–71, 2018.

[113] Sven Charleer, Joris Klerkx, Jose Luis Santos, and Erik Duval. Improving awareness and reflection through collaborative, interactive visualizations of badges. *AR-TEL@ EC-TEL*, 1103:69–81, 2013.

[114] Hui Chen and Kostadin Damevski. A teaching model for development of sensor-driven mobile applications. *ITICSE 2014 - Proceedings of the 2014 Innovation and Technology in Computer Science Education Conference*, pages 147–152, 2014.

[115] Zhengxing Chen, Christopher Amato, Truong-Huy D Nguyen, Seth Cooper, Yizhou Sun, and Magy Seif El-Nasr. Q-deckrec: A fast deck recommendation system for collectible card games. In *2018 IEEE Conference on Computational Intelligence and Games (CIG)*, pages 1–8. IEEE, 2018.

[116] Zhengxing Chen, Truong-Huy D Nguyen, Yuyu Xu, Christopher Amato, Seth Cooper, Yizhou Sun, and Magy Seif El-Nasr. The art of drafting: a team-oriented hero recommendation system for multiplayer online battle arena games. In *Proceedings of the 12th ACM Conference on Recommender Systems*, pages 200–208. ACM, 2018.

[117] Yuk Fai Cheong, Frank Pajares, and Paul S. Oberman. Motivation and academic help-seeking in high school computer science. *Computer Science Education*, 14(1):3–19, 2004.

[118] Donald Chinn and Tammy VanDeGrift. Gender and diversity in hiring software professionals. In *Proceeding of the fourth international workshop on Computing education research - ICER '08*, pages 39–50, New York, New York, USA, 2008. ACM Press.

[119] Luca Chittaro and Riccardo Sioni. Existential video games: Proposal and evaluation of an interactive reflection about death. *Entertainment computing*, 26:59–77, 2018.

[120] Alexander Cho, AM Tsaasan, and Constance Steinkuehler. The building blocks

of an educational esports league: lessons from year one in orange county high schools. In *Proceedings of the 14th International Conference on the Foundations of Digital Games*, pages 1–11, 2019.

[121] Chih-Yueh Chou, K Robert Lai, Po-Yao Chao, Chung Hsien Lan, and Tsung-Hsin Chen. Negotiation based adaptive learning sequences: Combining adaptivity and adaptability. *Computers & Education*, 88:215–226, 2015.

[122] Graig M Chow and Matteo Luzzeri. Post-event reflection: a tool to facilitate self-awareness, self-monitoring, and self-regulation in athletes. *Journal of Sport Psychology in Action*, 10(2):106–118, 2019.

[123] Anders Harboell Christiansen, Emil Gensby, and Bryan S Weber. Resolving simultaneity bias: Using features to estimate causal effects in competitive games. In *2019 IEEE Conference on Games (CoG)*, pages 1–8. IEEE, 2019.

[124] Cheng-Yu Chung and I-Han Hsiao. Investigating patterns of study persistence on self-assessment platform of programming problem-solving. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*, SIGCSE '20, page 162–168, New York, NY, USA, 2020. Association for Computing Machinery.

[125] Luke Church and Mariana Mărăşoiu. A fox not a hedgehog: What does ppig know? In *27th Workshop of the Psychology of Programming Interest Group - PPIG '16*, 2016.

[126] Geraldine Clarebout, Jan Elen, Lieve Luyten, and Hadewych Bamps. Assessing epistemological beliefs: Schommer's questionnaire revisited. *Educational Research and Evaluation*, 7(1):53–77, 2001.

[127] Dave Clarke, Tony Clear, Kathi Fisler, Matthias Hauswirth, Shriram Krishnamurthi, Joe Gibbs Politz, Ville Tirronen, and Tobias Wrigstad. In-Flow Peer Review. In *Proceedings of the Working Group Reports of the 2014 on Innovation

& Technology in Computer Science Education Conference - ITiCSE-WGR '14, pages 59–79, New York, New York, USA, 2014. ACM Press.

[128] Jody Clarke-Midura, Vicki Allan, and Kevin Close. Investigating the Role of Being a Mentor as a Way of Increasing Interest in CS. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education - SIGCSE '16*, pages 297–302, New York, New York, USA, Feb 2016. ACM Press.

[129] Alison. Clear, ACM Digital Library., and ACM. Special Interest Group on Computer Science Education. *Proceedings of the ninth annual international conference on International computing education research.* ACM, 2012.

[130] Tony Clear, JL Whalley, Phil Robbins, Anne Philpott, Anna Eckerdal, and Mikko-Jussi Laakso. Report on the final bracelet workshop: Auckland university of technology, september 2010. *Journal of Applied Computing and Information Technology*, 2011.

[131] Timothy J Cleary and Barry J Zimmerman. Self-regulation differences during athletic practice by experts, non-experts, and novices. *Journal of applied sport psychology*, 13(2):185–206, 2001.

[132] Timothy J Cleary, Barry J Zimmerman, and Tedd Keating. Training physical education students to self-regulate during basketball free throw practice. *Research Quarterly for Exercise and sport*, 77(2):251–262, 2006.

[133] Jacob Cohen. A coefficient of agreement for nominal scales. *Educational and psychological measurement*, 20(1):37–46, 1960.

[134] Peter A Cohen, James A. Kulik, and Chen-Lin C. Kulik. Educational outcomes of tutoring: A meta-analysis of findings. *American Educational Research Journal*, 19(2):237–248, 1982.

[135] Pamela M Cole, Sarah E Martin, and Tracy A Dennis. Emotion regulation as a sci-

entific construct: Methodological challenges and directions for child development research. *Child development*, 75(2):317–333, 2004.

[136] Kay Colthorpe, Jennifer Ogiji, Louise Ainscough, Kirsten Zimbardi, and Stephen Anderson. Effect of metacognitive prompts on undergraduate pharmacy students' self-regulated learning behavior. *American journal of pharmaceutical education*, 83(4), 2019.

[137] Mia Consalvo. *Cheating: Gaining advantage in videogames.* Mit Press, 2009.

[138] Linda Corrin and Paula De Barba. Exploring students' interpretation of feedback delivered through learning analytics dashboards. In *Proceedings of the ascilite 2014 conference*, pages 629–633. Dunedin, NZ, 2014.

[139] Savia Coutinho. Self-efficacy, metacognition, and performance. *North American Journal of Psychology*, 10(1), 2008.

[140] Michelle Craig, Diane Horton, Daniel Zingaro, and Danny Heap. Introducing and evaluating exam wrappers in cs2. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education*, SIGCSE '16, page 285–290, New York, NY, USA, 2016. Association for Computing Machinery.

[141] Michelle Craig, Andrew Petersen, and Jennifer Campbell. Answering the correct question. In *Proceedings of the ACM Conference on Global Computing Education*, CompEd '19, page 72–77, New York, NY, USA, 2019. Association for Computing Machinery.

[142] Kathryn Crawford and Alan Fekete. What do exam results really measure? In *Proceedings of the second Australasian conference on Computer science education - ACSE '97*, pages 185–190, New York, New York, USA, 1996. ACM Press.

[143] Kathryn Crawford and Alan Fekete. What do exam results really measure? In *Proceedings of the 2nd Australasian Conference on Computer Science Education*,

ACSE '97, page 185–190, New York, NY, USA, 1997. Association for Computing Machinery.

[144] Diana Cukierman, Donna McGee Thompson, Sheryl Guloy, Farimah Salimi, and Margarita Karpilovsky. Addressing challenges students face in first-year university Computing Science and Engineering Science courses. In *Proceedings of the Western Canadian Conference on Computing Education - WCCCE '14*, pages 1–5, New York, New York, USA, 2014. ACM Press.

[145] Diana R. Cukierman, Donna McGee Thompson, and Wayne Sun. The Academic Enhancement Program. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education - SIGCSE '19*, pages 686–692, New York, New York, USA, 2019. ACM Press.

[146] Quintin Cutts, Sarah Esper, Marlena Fecho, Stephen R. Foster, and Beth Simon. The abstraction transition taxonomy. In *Proceedings of the ninth annual international conference on International computing education research - ICER '12*, page 63, New York, New York, USA, 2012. ACM Press.

[147] Quintin Cutts, Charles Riedesel, Elizabeth Patitsas, Elizabeth Cole, Peter Donaldson, Bedour Alshaigy, Mirela Gutica, Arto Hellas, Edurne Larraza-Mendiluze, and Robert McCartney. Early Developmental Activities and Computing Proficiency. In *Proceedings of the 2017 ITiCSE Conference on Working Group Reports - ITiCSE-WGR '17*, volume 2018-Janua, pages 140–157, New York, New York, USA, 2017. ACM Press.

[148] John Dalbey and Marcia C Linn. The demands and requirements of computer programming: A literature review. *Journal of Educational Computing Research*, 1(3):253–274, 1985.

[149] Daniela Damian, Marian Petre, Mariel Miller, and Allyson F. Hadwin. Instructional strategies in the EGRET course. In *Proceedings of the Seventeenth Western*

Canadian Conference on Computing Education - WCCCE '12, page 41, New York, New York, USA, 2012. ACM Press.

[150] Daniela Damian, Marian Petre, Mariel Miller, and Allyson F. Hadwin. Instructional strategies in the egret course: An international graduate forum on becoming a researcher. *Proceedings of the 17th Western Canadian Conference on Computing Education, WCCCE 2012*, pages 41–45, 2012.

[151] Holger Danielsiek, Laura Toma, and Jan Vahrenhold. An Instrument to Assess Self-Efficacy in Introductory Algorithms Courses. In *Proceedings of the 2017 ACM Conference on International Computing Education Research - ICER '17*, volume 9, pages 217–225, New York, New York, USA, 2017. ACM Press.

[152] Pasqueline Dantas Scaico, Ruy José G.B. de Queiroz, and José Jorge Lima Dias. Analyzing How Interest in Learning Programming Changes During a CS0 Course. In *Proceedings of the 2017 ACM Conference on Innovation and Technology in Computer Science Education - ITiCSE '17*, volume Part F1286, pages 16–21, New York, New York, USA, 2017. ACM Press.

[153] James Davis, Michael Lachney, Zoe Zatz, William Babbitt, and Ron Eglash. A Cultural Computing Curriculum. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education - SIGCSE '19*, pages 1171–1175, New York, New York, USA, 2019. ACM Press.

[154] Giuseppe De Ruvo, Ewan Tempero, Andrew Luxton-Reilly, Gerard B. Rowe, and Nasser Giacaman. Understanding semantic style by analysing student code. In *Proceedings of the 20th Australasian Computing Education Conference on - ACE '18*, pages 73–82, New York, New York, USA, 2018. ACM Press.

[155] Jeanine A DeFalco, Jonathan P Rowe, Luc Paquette, Vasiliki Georgoulas-Sherry, Keith Brawner, Bradford W Mott, Ryan S Baker, and James C Lester. Detecting and addressing frustration in a serious game for military training. *International Journal of Artificial Intelligence in Education*, 28(2):152–193, 2018.

[156] Paul Denny, Brett A. Becker, Nigel Bosch, James Prather, Brent Reeves, and Jacqueline Whalley. Novice reflections during the transition to a new programming language. In *Proceedings of the 53rd ACM Technical Symposium V.1 on Computer Science Education*, SIGCSE 2022, page 948–954, New York, NY, USA, 2022. Association for Computing Machinery.

[157] Paul Denny, John Hamer, Andrew Luxton-Reilly, and Helen Purchase. Peerwise: Students sharing their multiple choice questions. In *Proceedings of the Fourth International Workshop on Computing Education Research*, pages 51–58, New York, NY, 2008. Association for Computing Machinery.

[158] Paul Denny, Andrew Luxton-Reilly, and John Hamer. The PeerWise System of Student Contributed Assessment Questions. Technical report.

[159] Paul Denny, Andrew Luxton-Reilly, John Hamer, Dana B. Dahlstrom, and Helen C. Purchase. Self-predicted and actual performance in an introductory programming course. In *Proceedings of the fifteenth annual conference on Innovation and technology in computer science education - ITiCSE '10*, page 118, New York, New York, USA, 2010. ACM Press.

[160] Paul Denny, James Prather, Brett A. Becker, Zachary Albrecht, Dastyni Loksa, and Raymond Pettit. A closer look at metacognitive scaffolding: Solving test cases before programming. In *Proceedings of the 19th Koli Calling International Conference on Computing Education Research*, Koli Calling '19, New York, NY, USA, 2019. Association for Computing Machinery.

[161] Paul Denny, Jacqueline Whalley, and Juho Leinonen. Promoting early engagement with programming assignments using scheduled automated feedback. In *Australasian Computing Education Conference*, pages 88–95, New York, NY, 2021. Association for Computing Machinery.

[162] Vania Dimitrova. Style-olm: Interactive open learner modelling. *International Journal of Artificial Intelligence in Education*, 13(1):35–78, 2003.

[163] Daniel L. Dinsmore, Patricia A. Alexander, and Sandra M. Loughlin. Focusing the conceptual lens on metacognition, self-regulation, and self-regulated learning. *Educational Psychology Review*, 20(4):391–409, Dec 2008.

[164] Scott Donaldson. Mechanics and metagame: Exploring binary expertise in league of legends. *Games and Culture*, 12(5):426–444, 2017.

[165] dotabuff. DOTABUFF. https://www.dotabuff.com/, 2021. Online; accessed 25 January 2021.

[166] Anders Drachen, Matthew Yancey, John Maguire, Derrek Chu, Iris Yuhui Wang, Tobias Mahlmann, Matthias Schubert, and Diego Klabajan. Skill-based differences in spatio-temporal team behaviour in defence of the ancients 2 (dota 2). In *2014 IEEE Games Media Entertainment*, pages 1–8. IEEE, 2014.

[167] Benedict Du Boulay. Some difficulties of learning to program. *Journal of Educational Computing Research*, 2(1):57–73, 1986.

[168] Benedict du Boulay, Tim O'Shea, and John Monk. The black box inside the glass box: presenting computing concepts to novices. *International Journal of Man-Machine Studies*, 14(3):237–249, 1981.

[169] Rachit Dubey, Pulkit Agrawal, Deepak Pathak, Thomas L Griffiths, and Alexei A Efros. Investigating human priors for playing video games. *arXiv preprint arXiv:1802.10217*, 2018.

[170] Rodrigo Duran, Jan Mikael Rybicki, Arto Hellas, and Sanna Suoranta. Towards a common instrument for measuring prior programming knowledge. In *Annual Conference on Innovation and Technology in Computer Science Education, ITiCSE*, pages 443–449. ACM, Jul 2019.

[171] Rodrigo Duran, Jan Mikael Rybicki, Juha Sorva, and Arto Hellas. Exploring the value of student self-evaluation in introductory programming. *ICER 2019 -*

Proceedings of the 2019 ACM Conference on International Computing Education Research, pages 121–130, 2019.

[172] Shree Durga, Magy Seif El-Nasr, Mariya Shiyko, Carmen Sceppa, Pamela Naab, and Lisa Andres. Leveraging play in health-based games to promote sustained behavior change in healthy eating and exercise. In *DiGRA Conference*. Citeseer, 2013.

[173] Shannon Duvall, Scott Spurlock, Dugald Ralph Hutchings, and Robert C. Duvall. Improving content learning and student perceptions in cs1 with scrumage. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education*, SIGCSE '21, page 474–480, New York, NY, USA, 2021. Association for Computing Machinery.

[174] Sebastian Dziallas and Sally Fincher. "I told you this last time, right?". In *Proceedings of the 2018 ACM Conference on International Computing Education Research - ICER '18*, pages 223–231, New York, New York, USA, 2018. ACM Press.

[175] Sebastian Dziallas, Sally Fincher, Colin G. Johnson, and Ian Utting. A first look at the year in computing. *Annual Conference on Innovation and Technology in Computer Science Education, ITiCSE*, Part F1286:275–280, 2017.

[176] J. Philip East. On models of and for teaching: Toward theory-based computing education. In *Proceedings of the 2006 international workshop on Computing education research - ICER '06*, volume 2006, page 41, New York, New York, USA, 2006. ACM Press.

[177] Joshua A Eaton, David J Mendonça, and Matthew-Donald D Sangster. Attack, damage and carry: Role familiarity and team performance in league of legends. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 62, pages 130–134. SAGE Publications Sage CA: Los Angeles, CA, 2018.

[178] Joshua A Eaton, Matthew-Donald D Sangster, Molly Renaud, David J Mendonca, and Wayne D Gray. Carrying the team: The importance of one player's survival for team success in league of legends. In *Proceedings of the human factors and ergonomics society annual meeting*, volume 61, pages 272–276. SAGE Publications Sage CA: Los Angeles, CA, 2017.

[179] Adam Eck, Leen-Kiat Soh, and Duane F. Shell. Investigating differences in wiki-based collaborative activities between student engagement profiles in cs1. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education*, SIGCSE '16, page 36–41, New York, NY, USA, 2016. ACM.

[180] Adam Eck, Leen-Kiat Soh, and Duane F. Shell. Investigating Differences in Wiki-based Collaborative Activities between Student Engagement Profiles in CS1. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education - SIGCSE '16*, pages 36–41, New York, New York, USA, 2016. ACM Press.

[181] Stephen H. Edwards, Joshua Martin, and Clfford A. Shaffer. Examining Classroom Interventions to Reduce Procrastination. In *Proceedings of the 2015 ACM Conference on Innovation and Technology in Computer Science Education - ITiCSE '15*, volume 2015-June, pages 254–259, New York, New York, USA, 2015. ACM Press.

[182] Anastasia Efklides. Metacognition: Defining its facets and levels of functioning in relation to self-regulation and co-regulation. *European Psychologist*, 13(4):277–287, 2008.

[183] Rylan Egan, Diana Cukierman, and Donna McGee Thompson. The academic enhancement program in introductory CS. In *Proceedings of the 16th annual joint conference on Innovation and technology in computer science education - ITiCSE '11*, page 278, New York, New York, USA, 2011. ACM Press.

[184] Markus Eger and Pablo Sauma Chacón. Deck archetype prediction in hearthstone.

In *International Conference on the Foundations of Digital Games*, pages 1–11, 2020.

[185] Magy Seif El-Nasr, Anders Drachen, and Alessandro Canossa. *Game analytics.* Springer, 2013.

[186] Magy Seif El-Nasr, Shree Durga, Mariya Shiyko, and Carmen Sceppa. Data-driven retrospective interviewing (ddri): a proposed methodology for formative evaluation of pervasive games. *Entertainment Computing*, 11:1–19, 2015.

[187] Magy Seif El-Nasr, Truong-Huy D Nguyen, Alessandro Canossa, and Anders Drachen. *Game Data Science.* Oxford University Press, 2021.

[188] Alexander Elias. *Goal setting in college athletics: The effects task oriented and ego-oriented goal setting strategies on physical endurance.* PhD thesis, Long Island University, The Brooklyn Center, 2013.

[189] Andrew J Elliot and Marcy A Church. A hierarchical model of approach and avoidance achievement motivation. *Journal of personality and social psychology*, 72(1):218, 1997.

[190] Andrew J Elliot and Holly A McGregor. A 2× 2 achievement goal framework. *Journal of personality and social psychology*, 80(3):501, 2001.

[191] Alex Endert, M Shahriar Hossain, Naren Ramakrishnan, Chris North, Patrick Fiaux, and Christopher Andrews. The human is the loop: new directions for visual analytics. *Journal of intelligent information systems*, 43(3):411–435, 2014.

[192] Daniel Epstein, Felicia Cordeiro, Elizabeth Bales, James Fogarty, and Sean Munson. Taming data complexity in lifelogs: exploring visual cuts of personal informatics data. In *Proceedings of the 2014 conference on Designing interactive systems*, pages 667–676, 2014.

[193] Ali Erkan, John Barr, Tony Clear, Cruz Izu, Cristian Jose Lopez del Alamo, Hanan Mohammed, and Mahadev Nadimpalli. Developing a Holistic Understanding of

Systems and Algorithms through Research Papers. In *Proceedings of the 2017 ITiCSE Conference on Working Group Reports - ITiCSE-WGR '17*, volume 2018-Janua, pages 86–104, New York, New York, USA, 2017. ACM Press.

[194] Peggy A. Ertmer and Timothy J. Newby. Expert learner: Strategic, self-regulated, and reflective. *Instructional Science*, 24(1):1–24, 1996.

[195] esports.net. Is esports a sport? countries where esports is considered a sport. 2021.

[196] Jose Esteves, Konstantina Valogianni, and Anita Greenhill. Online social games: The effect of social comparison elements on continuance behaviour. *Information & Management*, 58(4):103452, 2021.

[197] Anneli Eteläpelto. Metacognition and the expertise of computer program comprehension. *Scandinavian Journal of Educational Research*, 37(3):243–254, 1993.

[198] MD Evans. A new emphasis & pedagogy for a cs1 course. *ACM SIGCSE Bulletin*, 28(3):12–16, 1996.

[199] Marisa Exter. Comparing educational experiences and on-the-job needs of educational software designers. *SIGCSE 2014 - Proceedings of the 45th ACM Technical Symposium on Computer Science Education*, pages 355–360, 2014.

[200] Muhammad Fahmi, Dedi Purwana, et al. The effect of coach leadership, team cooperation, and achievement motivation on the performance of the depok city athlete. *IJHCM (International Journal of Human Capital Management)*, 3(1):27–38, 2019.

[201] Katrina Falkner, Nickolas J.G. Falkner, and Rebecca Vivian. Collaborative learning and anxiety. In *Proceeding of the 44th ACM technical symposium on Computer science education - SIGCSE '13*, page 227, New York, New York, USA, 2013. ACM Press.

[202] Katrina Falkner, Claudia Szabo, Rebecca Vivian, and Nickolas Falkner. Evolution of software development strategies. In *Proceedings of the 37th International Conference on Software Engineering - Volume 2*, ICSE '15, page 243–252, Piscataway, NJ, 2015. IEEE Press.

[203] Katrina Falkner, Rebecca Vivian, and Nickolas J.G. Falkner. Identifying computer science self-regulated learning strategies. In *Proceedings of the 2014 Conference on Innovation and Technology in Computer Science Education*, ITiCSE '14, page 291–296, New York, NY, USA, 2014. ACM.

[204] Nickolas Falkner, Claudia Szabo, and Antti Knutas. Classifying the nature of lecturer intervention on computer science student behaviour. In *Proceedings of the 17th Koli Calling Conference on Computing Education Research - Koli Calling '17*, pages 128–132, New York, New York, USA, 2017. ACM Press.

[205] Nickolas Falkner, Rebecca Vivian, David Piper, and Katrina Falkner. Increasing the effectiveness of automated assessment by increasing marking granularity and feedback units. In *Proceedings of the 45th ACM Technical Symposium on Computer Science Education*, SIGCSE '14, page 9–14, New York, NY, USA, 2014. Association for Computing Machinery.

[206] Nickolas J.G. Falkner and Katrina E. Falkner. A fast measure for identifying at-risk students in computer science. In *Proceedings of the ninth annual international conference on International computing education research - ICER '12*, page 55, New York, New York, USA, 2012. ACM Press.

[207] Joey R Fanfarelli. Expertise in professional overwatch play. *International Journal of Gaming and Computer-Mediated Simulations (IJGCMS)*, 10(1):1–22, 2018.

[208] Alan Fekete and Antony Greening. Designing closed laboratories for a computer science course. In *Proceedings of the twenty-seventh SIGCSE technical symposium on Computer science education - SIGCSE '96*, pages 295–299, New York, New York, USA, 1996. ACM Press.

[209] Alan Fekete, Judy Kay, Jeff Kingston, and Kapila Wimalaratne. Supporting reflection in introductory computer science. In *Proceedings of the thirty-first SIGCSE technical symposium on Computer science education - SIGCSE '00*, volume 32, pages 144–148, New York, New York, USA, Mar 2000. ACM Press.

[210] Russell Feldhausen, Joshua Levi Weese, and Nathan H. Bean. Increasing Student Self-Efficacy in Computational Thinking via STEM Outreach Programs. In *Proceedings of the 49th ACM Technical Symposium on Computer Science Education - SIGCSE '18*, volume 2018-Janua, pages 302–307, New York, New York, USA, 2018. ACM Press.

[211] Chia-Yen Feng and Ming-Puu Chen. The effects of goal specificity and scaffolding on programming performance and self-regulation in game design. *British Journal of Educational Technology*, 45(2):285–302, 2014.

[212] Andrea Ferrario, Michele Loi, and Eleonora Viganò. In ai we trust incrementally: a multi-layer model of trust to analyze human-artificial intelligence interactions. *Philosophy & Technology*, pages 1–17, 2019.

[213] Ana Figueiras. Towards the understanding of interaction in information visualization. In *2015 19th International Conference on Information Visualisation*, pages 140–147. IEEE, 2015.

[214] Logan Fiorella and Richard E Mayer. based aids for learning with a computer-based game. *Journal of Educational Psychology*, 104(4):1074, 2012.

[215] Kasper Fisker, Davin McCall, Michael Kölling, and Bruce Quig. Group work support for the bluej ide. In *Proceedings of the 13th Annual Conference on Innovation and Technology in Computer Science Education*, pages 163–168, New York, NY, 2008. Association for Computing Machinery.

[216] Abraham E. Flanigan, Markeya S. Peteranetz, Duane F. Shell, and Leen-Kiat Soh. Exploring Changes in Computer Science Students' Implicit Theories of Intelligence

Across the Semester. In *Proceedings of the eleventh annual International Conference on International Computing Education Research - ICER '15*, pages 161–168, New York, New York, USA, 2015. ACM Press.

[217] J. H. Flavell. Metacognitive aspects of problem solving. *The Nature of Intelligence*, pages 231–235, 1976.

[218] John H Flavell. Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34(10):906, 1979.

[219] Rowanne Fleck and Geraldine Fitzpatrick. Reflecting on reflection: framing a design landscape. In *Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction*, pages 216–223, 2010.

[220] Rachel Flynn. Acute effects of physically active versus inactive video game play on executive functioning skills in children. In *Proceedings of the International Conference on the Foundations of Digital Games - FDG '12*, page 264, New York, New York, USA, 2012. ACM Press.

[221] Jessica Formosa, Nicholas O'Donnell, Ella M Horton, Selen Türkay, Regan L Mandryk, Michael Hawks, and Daniel Johnson. Definitions of esports: A systematic review and thematic analysis. *Proceedings of the ACM on Human-Computer Interaction*, 6(CHI PLAY):1–45, 2022.

[222] Allan Fowler, Johanna Pirker, Ian Pollock, Bruno Campagnola De Paula, Maria Emilia Echeveste, and Marcos J. Gómez. Understanding the benefits of game jams: Exploring the potential for engaging young learners in STEM. *Proceedings of the 2016 ITiCSE Working Group Reports, ITiCSE 2016*, pages 119–135, 2016.

[223] Emily Fox and Michelle Riconscente. Metacognition and self-regulation in james, piaget, and vygotsky. *Educational Psychology Review*, 20(4):373–389, 2008.

[224] Diana Franklin, Jean Salac, Zachary Crenshaw, Saranya Turimella, Zipporah Klain, Marco Anaya, and Cathy Thomas. Exploring student behavior using the tipp&see learning strategy. In *Proceedings of the 2020 ACM Conference on International Computing Education Research*, ICER '20, page 91–101, New York, NY, USA, 2020. Association for Computing Machinery.

[225] Guo Freeman and Donghee Yvette Wohn. Social support in esports: building emotional and esteem support from instrumental support interactions in a highly competitive environment. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, pages 435–447, 2017.

[226] Guo Freeman and Donghee Yvette Wohn. Understanding esports team formation and coordination. *Computer supported cooperative work (CSCW)*, 28(1):95–126, 2019.

[227] Stephen Frezza, Mats Daniels, Arnold Pears, Åsa Cajander, Viggo Kann, Amanpreet Kapoor, Roger McDermott, Anne Kathrin Peters, Mihaela Sabin, and Charles Wallace. Modelling competencies for computing education beyond 2020: A research based approach to defining competencies in the computing disciplines. *Annual Conference on Innovation and Technology in Computer Science Education, ITiCSE*, pages 148–174, 2018.

[228] Fridai. Fridai Voice Assistant. https://hellofridai.com/, 2021. Online; accessed December 9th 2021.

[229] John Fritz. Classroom walls that talk: Using online course activity data of successful students to raise self-awareness of underperforming peers. *The Internet and Higher Education*, 14(2):89–97, 2011.

[230] Ursula Fuller, Charles Riedesel, Errol Thompson, Colin G. Johnson, Tuukka Ahoniemi, Diana Cukierman, Isidoro Hernán-Losada, Jana Jackova, Essi Lahtinen, Tracy L. Lewis, and Donna McGee Thompson. Developing a computer

science-specific learning taxonomy. In *Working group reports on ITiCSE on Innovation and technology in computer science education - ITiCSE-WGR '07*, page 152, New York, New York, USA, 2007. ACM Press.

[231] Shelly L Gable and Harry T Reis. Good news! capitalizing on positive events in an interpersonal context. In *Advances in experimental social psychology*, volume 42, pages 195–257. Elsevier, 2010.

[232] André R Gagné, Magy Seif El-Nasr, and Chris D Shaw. A deeper look at the use of telemetry for analysis of player behavior in rts games. In *International Conference on Entertainment Computing*, pages 247–257. Springer, 2011.

[233] Alex Games and Luke Kane. Exploring adolescent's STEM learning through scaffolded game design. In *Proceedings of the 6th International Conference on Foundations of Digital Games - FDG '11*, pages 1–8, New York, New York, USA, 2011. ACM Press.

[234] Rita Garcia, Katrina Falkner, and Rebecca Vivian. Instructional Framework for CS1 Question Activities. In *Proceedings of the 2019 ACM Conference on Innovation and Technology in Computer Science Education - ITiCSE '19*, pages 189–195, New York, New York, USA, 2019. ACM Press.

[235] D Randy Garrison and Zehra Akyol. Thinking collaboratively in educational environments: Shared metacognition and co-regulation in communities of inquiry. In *Educational developments, practices and effectiveness*, pages 39–52. Springer, Cham, CH, 2015.

[236] Ann Q Gates. Integrating a problem-solving methodology and group skills into cs1. In *Proceedings of 9th Conference on Software Engineering Education*, pages 6–15, Piscataway, NJ, 1996. IEEE.

[237] Helen Gavin. Thematic analysis. *Understanding research methods and statistics in psychology*, pages 273–282, 2008.

[238] James Paul Gee. What video games have to teach us about learning and literacy. *Computers in entertainment (CIE)*, 1(1):20–20, 2003.

[239] James Paul Gee. Learning by design: Games as learning machines. *Interactive educational multimedia: IEM*, pages 15–23, 2004.

[240] Pauline Gibbons. *Scaffolding language, scaffolding learning.* Heinemann Portsmouth, NH, 2002.

[241] David Ginat. Metacognitive awareness utilized for learning control elements in algorithmic problem solving. *ACM SIGCSE Bulletin*, 33(3):81–84, Sep 2001.

[242] David Ginat. The greedy trap and learning from mistakes. *SIGCSE Bulletin (ACM, Special Interest Group on Computer Science Education)*, pages 11–15, 2003.

[243] David Ginat and Ronit Shmalo. Constructive use of errors in teaching CS1. In *Proceeding of the 44th ACM technical symposium on Computer science education - SIGCSE '13*, page 353, New York, New York, USA, 2013. ACM Press.

[244] Daniela Giordano, Francesco Maiorana, Andrew Paul Csizmadia, Simon Marsden, Charles Riedesel, Shitanshu Mishra, and Lina Vinikienė. New Horizons in the Assessment of Computer Science at School and Beyond. In *Proceedings of the 2015 ITiCSE on Working Group Reports - ITICSE-WGR '15*, pages 117–147, New York, New York, USA, 2015. ACM Press.

[245] Lisa Given. *100 Questions (and Answers) About Qualitative Research*, volume 68. SAGE Publications, Thousand Oaks, CA, 2016.

[246] Barney G. Glaser and Anselm L. Strauss. *The Discovery of Grounded Theory: Strategies for Qualitative Research.* Aldine de Gruyter, NY, NY, 1967.

[247] Jennifer L Goetz, Dacher Keltner, and Emiliana Simon-Thomas. Compassion: an evolutionary analysis and empirical review. *Psychological bulletin*, 136(3):351, 2010.

[248] Michael Goldweber, John Impagliazzo, Iouri A. Bogoiavlenski, A. G. Clear, Gordon Davies, Hans Flack, J. Paul Myers, and Richard Rasala. Historical perspectives on the computing curriculum (report of the ITiCSE '97 working group on historical perspectives in computing education). In *The supplemental Proceedings of the conference on Integrating technology into computer science education: working group reports and supplemental proceedings - ITiCSE-WGR '97*, volume Part F1293, pages 94–111, New York, New York, USA, 1997. ACM Press.

[249] Anabela Jesus Gomes, Álvaro Nuno Santos, and António José Mendes. A study on students' behaviours and attitudes towards learning to program. *Annual Conference on Innovation and Technology in Computer Science Education, ITiCSE*, pages 132–137, 2012.

[250] Marcos J. Gomez, Marco Moresi, and Luciana Benotti. Text-based Programming in Elementary School. In *Proceedings of the 2019 ACM Conference on Innovation and Technology in Computer Science Education - ITiCSE '19*, pages 402–408, New York, New York, USA, 2019. ACM Press.

[251] Tiago Gonçalves, Pedro Vieira, Ana Paula Afonso, Maria Beatriz Carmo, and Tiago Moucho. Analysing player performance with animated maps. In *2018 22nd international conference information visualisation (IV)*, pages 103–109. IEEE, 2018.

[252] Graciela Gonzalez. A systematic approach to active and cooperative learning in cs1 and its effects on cs2. In *Proceedings of the 37th SIGCSE technical symposium on Computer science education*, pages 133–137, New York, NY, 2006. Association for Computing Machinery.

[253] Joanna Goode, Jane Margolis, and Gail Chapman. Curriculum is not enough. In *Proceedings of the 45th ACM technical symposium on Computer science education - SIGCSE '14*, pages 493–498, New York, New York, USA, 2014. ACM Press.

[254] Jamie Gorson and Eleanor O'Rourke. How Do Students Talk About Intelligence? In *Proceedings of the 2019 ACM Conference on International Computing Education Research - ICER '19*, pages 21–29, New York, New York, USA, 2019. ACM Press.

[255] Jamie Gorson and Eleanor O'Rourke. Why do cs1 students think they're bad at programming? investigating self-efficacy and self-assessments at three universities. In *Proceedings of the 2020 ACM Conference on International Computing Education Research*, ICER '20, page 170–181, New York, NY, USA, 2020. Association for Computing Machinery.

[256] GOSU.AI. GOSU Voice Assistant. https://gosu.ai/, 2021. Online; accessed April 15th 2021.

[257] Don Gotterbarn. The ethical computer practitioner—licensing the moral community. *ACM SIGCSE Bulletin*, 30(2):8–10, Jun 1998.

[258] Sten Govaerts, Katrien Verbert, Erik Duval, and Abelardo Pardo. The student activity meter for awareness and self-reflection. In *CHI'12 Extended Abstracts on Human Factors in Computing Systems*, pages 869–884. 2012.

[259] Steve Graham, Charles Macarthur, Shirley Schwartz, and Victoria Page-Voth. Improving the compositions of students with learning disabilities using a strategy involving product and process goal setting. *Exceptional Children*, 58(4):322–334, 1992.

[260] Simon Gray, Caroline St. Clair, Richard James, and Jerry Mead. Suggestions for graduated exposure to programming concepts using fading worked examples. In *Proceedings of the third international workshop on Computing education research - ICER '07*, page 99, New York, New York, USA, 2007. ACM Press.

[261] Tony Greening, Judy Kay, Jeffrey H Kingston, and Kathryn Crawford. Results of a PBL trial in first-year computer science. In *Proceedings of the second Australasian*

conference on Computer science education - ACSE '97, volume Part F1293, pages 201–206, New York, New York, USA, 1996. ACM Press.

[262] James J Gross. The emerging field of emotion regulation: An integrative review. *Review of general psychology*, 2(3):271–299, 1998.

[263] James J Gross. Emotion regulation: Past, present, future. *Cognition & emotion*, 13(5):551–573, 1999.

[264] James J Gross. Emotion regulation: conceptual and empirical foundations. 2014.

[265] James J Gross and Ross A Thompson. Emotion regulation: Conceptual foundations. inj. j. gross (ed.), handbook of emotion regulation. *New York: The Guillford Press*, 3:25, 2007.

[266] Charlotte N Gunawardena, Mary Beth Hermans, Damien Sanchez, Carol Richmond, Maribeth Bohley, and Rebekah Tuttle. A theoretical framework for building online communities of practice with social networking tools. *Educational media international*, 46(1):3–16, 2009.

[267] David Gunning. Explainable artificial intelligence (xai). *Defense advanced research projects agency (DARPA), nd Web*, 2(2):1, 2017.

[268] Edwin Ray Guthrie. Conditioning as a principle of learning. *Psychological review*, 37(5):412, 1930.

[269] Mark Guzdial, Shriram Krishnamurthi, Juha Sorva, and Jan Vahrenhold. Notional Machines and Programming Language Semantics in Education (Dagstuhl Seminar 19281). *Dagstuhl Reports*, 9(7):1–23, 2019.

[270] Steve Hadfield, Troy Weingart, Joel Coffman, David Caswell, Barry Fagin, Traci Sarmiento, and Paul Graham. Streamlining Computer Science Curriculum Development and Assessment using the New ABET Student Outcomes. In *Proceedings of the Western Canadian Conference on Computing Education - WCCCE '19*, pages 1–6, New York, New York, USA, 2019. ACM Press.

[271] Paulina Haduong and Karen Brennan. Helping K-12 teachers get unstuck with scratch: The design of an online professional learning experience. *SIGCSE 2019 - Proceedings of the 50th ACM Technical Symposium on Computer Science Education*, pages 1095–1101, 2019.

[272] Allyson Hadwin and Mika Oshige. Self-regulation, coregulation, and socially shared regulation: Exploring perspectives of social in self-regulated learning theory. *Teachers College Record*, 113(2):240–264, 2011.

[273] Allyson Fiona Hadwin, Sanna Järvelä, and Mariel Miller. Self-regulated, co-regulated, and socially shared regulation of learning. *Handbook of self-regulation of learning and performance*, 30:65–84, 2011.

[274] Nour Halabi, Günter Wallner, and Pejman Mirza-Babaei. Assessing the impact of visual design on the interpretation of aggregated playtesting data visualization. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, pages 639–650, 2019.

[275] Erwin Handoko, Susie L Gronseth, Sara G McNeil, Curtis J Bonk, and Bernard R Robin. Goal setting and mooc completion: A study on the role of self-regulated learning in student performance in massive open online courses. *International Review of Research in Open and Distributed Learning*, 20(3), 2019.

[276] Danita Hartley and Antonija Mitrovic. Supporting learning by opening the student model. In *International Conference on Intelligent Tutoring Systems*, pages 453–462. Springer, 2002.

[277] Uri Hasson, Janice Chen, and Christopher J Honey. Hierarchical process memory: memory as an integral component of information processing. *Trends in cognitive sciences*, 19(6):304–313, 2015.

[278] Neville Hatton and David Smith. Reflection in teacher education: Towards definition and implementation. *Teaching and teacher education*, 11(1):33–49, 1995.

[279] Neville Hatton and David W Smith. Reflection in teacher education: Towards definition and implementation. volume 11, pages 33–49, 1995.

[280] Erik Hazzard. Data visualization in games, 2014.

[281] Huan He, Bo Dong, Qinghua Zheng, and Guobin Li. VUC: Visualizing Daily Video Utilization to Promote Student Engagement in Online Distance Education. In *Proceedings of the ACM Conference on Global Computing Education - CompEd '19*, pages 99–105, New York, New York, USA, May 2019. ACM Press.

[282] Arto Hellas, Soohyun Nam Liao, Petri Ihantola, Andrew Petersen, Vangel V. Ajanovski, Mirela Gutica, Timo Hynninen, Antti Knutas, Juho Leinonen, and Chris Messom. Predicting academic performance: a systematic literature review. In *Proceedings Companion of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education - ITiCSE 2018 Companion*, pages 175–199, New York, New York, USA, 2018. ACM Press.

[283] Werner F Helsen, Janet L Starkes, and Nicola J Hodges. Team sports and the theory of deliberate practice. *Journal of Sport and Exercise psychology*, 20(1):12–34, 1998.

[284] Dean Hendrix, Lakshman Myneni, Hari Narayanan, and Margaret Ross. Implementing studio-based learning in cs2. In *Proceedings of the 41st ACM Technical Symposium on Computer Science Education*, SIGCSE '10, page 505–509, New York, NY, USA, 2010. ACM.

[285] Dean Hendrix, Lakshman Myneni, Hari Narayanan, and Margaret Ross. Implementing studio-based learning in CS2. In *Proceedings of the 41st ACM technical symposium on Computer science education - SIGCSE '10*, page 505, New York, New York, USA, 2010. ACM Press.

[286] Geoffrey L Herman, Lisa Kaczmarczyk, Michael C Loui, and Craig Zilles. Proof by incomplete enumeration and other logical misconceptions. In *Proceeding of*

the fourth international workshop on Computing education research - ICER '08, pages 59–70, New York, New York, USA, 2008. ACM Press.

[287] Joseph Hesketh, Christoph Sebastian Deterding, and Jeremy Gow. How players learn team-versus-team esports: First results from a grounded theory study. In *DiGRA'20 Abstract-Proceedings of the 2020 DiGRA International Conference*. York, 2020.

[288] Andrew D. Hilton, Genevieve M. Lipp, and Susan H. Rodger. Translation from Problem to Code in Seven Steps. In *Proceedings of the ACM Conference on Global Computing Education - CompEd '19*, pages 78–84, New York, New York, USA, May 2019. ACM Press.

[289] Ivo van Hilvoorde and Niek Pot. Embodiment and fundamental motor skills in esports. *Sport, Ethics and Philosophy*, 10(1):14–27, 2016.

[290] Mark E. Hoffman, Paul V. Anderson, and Magnus Gustafsson. Workplace Scenarios to Integrate Communication Skills and Content: A Case Study. *SIGCSE 2014 - Proceedings of the 45th ACM Technical Symposium on Computer Science Education*, pages 349–354, 2014.

[291] Caitlin Holman, Stephen Aguilar, and Barry Fishman. Gradecraft: What can we learn from a game-inspired learning management system? In *Proceedings of the third international conference on learning analytics and knowledge*, pages 260–264, 2013.

[292] Hilary J Holz, Anne Applin, Bruria Haberman, Donald Joyce, Helen Purchase, and Catherine Reed. Research methods in computing. In *Working group reports on ITiCSE on Innovation and technology in computer science education - ITiCSE-WGR '06*, page 96, New York, New York, USA, 2006. ACM Press.

[293] Hee Jung Hong. esports: the need for a structured support system for players. *European Sport Management Quarterly*, pages 1–24, 2022.

[294] Danial Hooshyar, Margus Pedaste, Katrin Saks, Äli Leijen, Emanuele Bardone, and Minhong Wang. Open learner models in supporting self-regulated learning in higher education: A systematic literature review. *Computers & Education*, 154:103878, 2020.

[295] Danial Hooshyar, Margus Pedaste, Yeongwook Yang, Liina Malva, Gwo-Jen Hwang, Minhong Wang, Heuiseok Lim, and Dejan Delev. From gaming to computational thinking: An adaptive educational computer game-based learning approach. *Journal of Educational Computing Research*, 59(3):383–409, 2021.

[296] Danial Hooshyar, Moslem Yousefi, and Heuiseok Lim. A systematic review of data-driven approaches in player modeling of educational games. *Artificial Intelligence Review*, 52(3):1997–2017, 2019.

[297] Britton Horn, Amy K Hoover, Jackie Barnes, Yetunde Folajimi, Gillian Smith, and Casper Harteveld. Opening the black box of play: Strategy analysis of an educational game. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play*, pages 142–153. ACM, 2016.

[298] Diane Horton, Jennifer Campbell, and Michelle Craig. Online CS1: Who Enrols, why, and how do they do? In *SIGCSE 2016 - Proceedings of the 47th ACM Technical Symposium on Computing Science Education*, pages 335–340. ACM, Inc, Feb 2016.

[299] I-H Hsiao, Sergey Sosnovsky, and Peter Brusilovsky. Guiding students to the right questions: adaptive navigation support in an e-learning system for java programming. *Journal of Computer Assisted Learning*, 26(4):270–283, 2010.

[300] I-Han Hsiao and Peter Brusilovsky. Guiding and motivating students through open social student modeling: Lessons learned. *Teachers College Record*, 119(3), 2017.

[301] I-Han Hsiao, Julio Guerra, Denis Parra, Fedor Bakalov, Birgitta König-Ries, and

Peter Brusilovsky. Comparative social visualization for personalized e-learning. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*, pages 303–307, 2012.

[302] Dandan Huang, Melanie Tory, Bon Adriel Aseniero, Lyn Bartram, Scott Bateman, Sheelagh Carpendale, Anthony Tang, and Robert Woodbury. Personal visualization and personal visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 21(3):420–433, 2014.

[303] Eddie Huang, Hannah Valdiviejas, and Nigel Bosch. I'm sure! Automatic detection of metacognition in online course discussion forums. In *Proceedings of the 8th International Conference on Affective Computing and Intelligent Interaction (ACII 2019)*, pages 241–247, Piscataway, NJ, 2019. IEEE.

[304] Alison Hull. *Motivational and Metacognitive Feedback in an ITS: Linking Past States and Experiences to Current Problems.* PhD thesis, University of Sussex, 2014.

[305] Alison Hull and Benedict du Boulay. Motivational and metacognitive feedback in sql-tutor*. *Computer Science Education*, 25(2):238–256, 2015.

[306] Christopher Hundhausen, Anukrati Agrawal, Dana Fairbrother, and Michael Trevisan. Does studio-based instruction work in CS 1? In *Proceedings of the 41st ACM technical symposium on Computer science education - SIGCSE '10*, page 500, New York, New York, USA, 2010. ACM Press.

[307] Christopher Hundhausen, Anukrati Agrawal, Dana Fairbrother, and Michael Trevisan. Does studio-based instruction work in cs 1? an empirical comparison with a traditional approach. In *Proceedings of the 41st ACM Technical Symposium on Computer Science Education*, SIGCSE '10, page 500–504, New York, NY, USA, 2010. ACM.

[308] Christopher D. Hundhausen, Pawan Agarwal, and Michael Trevisan. Online vs.

face-to-face pedagogical code reviews. In *Proceedings of the 42nd ACM technical symposium on Computer science education - SIGCSE '11*, page 117, New York, New York, USA, 2011. ACM Press.

[309] Robin Hunicke. The case for dynamic difficulty adjustment in games. In *Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology*, pages 429–433, 2005.

[310] Tarja-Riitta Hurme, Sanna Järvelä, Kaarina Merenluoto, and Pekka Salonen. What makes metacognition as socially shared in mathematical problem solving? In *Metacognition: Fundaments, Applications, and Trends*, pages 259–276. Springer, Cham, CH, 2015.

[311] Ioanna Iacovides, Anna L Cox, Ara Avakian, and Thomas Knoll. Player strategies: Achieving breakthroughs and progressing in single-player and cooperative games. In *Proceedings of the first ACM SIGCHI annual symposium on Computer-human interaction in play*, pages 131–140, 2014.

[312] Ioanna Iacovides, Anna L Cox, and Thomas Knoll. Learning the game: breakdowns, breakthroughs and player strategies. In *CHI'14 Extended Abstracts on Human Factors in Computing Systems*, pages 2215–2220. 2014.

[313] Ioanna Iacovides, Anna L Cox, Patrick McAndrew, James Aczel, and Eileen Scanlon. Game-play breakdowns and breakthroughs: exploring the relationship between action, understanding, and involvement. *Human–computer interaction*, 30(3-4):202–231, 2015.

[314] Dirk Ifenthaler. Determining the effectiveness of prompts for self-regulated learning in problem-solving scenarios. *Journal of Educational Technology & Society*, 15(1):38–52, 2012.

[315] Kalle Ilves, Juho Leinonen, and Arto Hellas. Supporting self-regulated learning with visualizations in online learning environments. In *Proceedings of the 49th*

ACM Technical Symposium on Computer Science Education, SIGCSE '18, page 257–262, New York, NY, USA, 2018. Association for Computing Machinery.

[316] Ville Isomöttönen and Aletta Nylén. How students get going: triggers for students' learning in project-based education. In *Proceedings of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education - ITiCSE 2018*, pages 117–122, New York, New York, USA, 2018. ACM Press.

[317] Ville Isomöttönen and Ville Tirronen. Flipping and blending—an action research project on improving a functional programming course. *ACM Trans. Comput. Educ.*, 17(1), September 2016.

[318] Cruz Izu, Cheryl Pope, and Amali Weerasinghe. On the Ability to Reason About Program Behaviour. In *Proceedings of the 2017 ACM Conference on Innovation and Technology in Computer Science Education - ITiCSE '17*, volume Part F1286, pages 305–310, New York, New York, USA, 2017. ACM Press.

[319] Andrzej Janusz, Dominik Slezak, Sebastian Stawicki, and Krzysztof Stencel. Sensei: an intelligent advisory system for the esport community and casual players. In *2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI)*, pages 754–757. IEEE, 2018.

[320] Nithesh Javvaji, Casper Harteveld, and Magy Seif El-Nasr. Understanding player patterns by combining knowledge-based data abstraction with interactive visualization. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, 2020.

[321] Chaima Jemmali. May's journey: A serious game to teach middle and high school girls programming. 2016.

[322] Chaima Jemmali, Erica Kleinman, Sara Bunian, Mia Victoria Almeda, Elizabeth Rowe, and Magy Seif El-Nasr. Maads: Mixed-methods approach for the analysis

of debugging sequences of beginner programmers. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*, pages 86–92, 2020.

[323] Maria Kallia and Sue Sentance. Learning to Use Functions: The Relationship Between Misconceptions and Self-Efficacy. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education - SIGCSE '19*, pages 752–758, New York, New York, USA, 2019. ACM Press.

[324] Shekhar Kalra, Charles Thevathayan, and Margaret Hamilton. Developing industry-relevant higher order thinking skills in computing students. In *Proceedings of the 2020 ACM Conference on Innovation and Technology in Computer Science Education*, ITiCSE '20, page 294–299, New York, NY, USA, 2020. Association for Computing Machinery.

[325] Slava Kalyuga. The expertise reversal effect. In *Managing cognitive load in adaptive multimedia learning*, pages 58–80. IGI Global, 2009.

[326] Jimoon Kang, June Seop Yoon, and Byungjoo Lee. How ai-based training affected the performance of professional go players. In *CHI Conference on Human Factors in Computing Systems*, pages 1–12, 2022.

[327] Viggo Kann and Anna-Karin Högfeldt. Effects of a program integrating course for students of computer science and engineering. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education*, SIGCSE '16, page 510–515, New York, NY, USA, 2016. Association for Computing Machinery.

[328] Avi Kaplan and Martin L Maehr. The contributions and prospects of goal orientation theory. *Educational Psychology Review*, 19(2):141–184, 2007.

[329] Amanpreet Kapoor and Christina Gardner-McCune. Exploring the participation of cs undergraduate students in industry internships. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*, SIGCSE '20, page 1103–1109, New York, NY, USA, 2020. Association for Computing Machinery.

[330] Michael Karlsson. Information visualisation in games: How do you know what you know?, 2016.

[331] Ioannis Karvelas, Annie Li, and Brett A. Becker. The effects of compilation mechanisms and error message presentation on novice programmer behavior. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*, SIGCSE '20, page 759–765, New York, NY, USA, 2020. Association for Computing Machinery.

[332] Douglas F Kauffman. Self-regulated learning in web-based environments: Instructional tools designed to facilitate cognitive strategy use, metacognitive processing, and motivational beliefs. *Journal of educational computing research*, 30(1-2):139–161, 2004.

[333] Douglas F Kauffman, Ruomeng Zhao, and Ya-Shu Yang. Effects of online note taking formats and self-monitoring prompts on learning from online text: Using technology to enhance self-regulated learning. *Contemporary Educational Psychology*, 36(4):313–322, 2011.

[334] Alice Kerly, Richard Ellis, and Susan Bull. Calmsystem: a conversational agent for learner modelling. In *International conference on innovative techniques and applications of artificial intelligence*, pages 89–102. Springer, 2007.

[335] Ada S. Kim and Amy J. Ko. A pedagogical analysis of online coding tutorials. In *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education*, SIGCSE '17, page 321–326, New York, NY, USA, 2017. ACM.

[336] Ada S. Kim and Amy J. Ko. A pedagogical analysis of online coding tutorials. In *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education*, SIGCSE '17, page 321–326, New York, NY, USA, 2017. Association for Computing Machinery.

[337] Yea-Seul Kim, Katharina Reinecke, and Jessica Hullman. Explaining the gap:

Visualizing one's predictions improves recall and comprehension of data. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 1375–1386, 2017.

[338] Päivi Kinnunen and Lauri Malmi. Why students drop out CS1 course? *ICER 2006 - Proceedings of the 2nd International Computing Education Research Workshop*, 2006:97–108, 2006.

[339] Päivi Kinnunen, Veijo Meisalo, and Lauri Malmi. Have we missed something? In *Proceedings of the Sixth international workshop on Computing education research - ICER '10*, page 13, New York, New York, USA, 2010. ACM Press.

[340] Päivi Kinnunen and Beth Simon. CS majors' self-efficacy perceptions in CS1: Results in light of social cognitive theory. *ICER'11 - Proceedings of the ACM SIGCSE 2011 International Computing Education Research Workshop*, pages 19–26, 2011.

[341] Michael S. Kirkpatrick, Mohamed Aboutabl, David Bernstein, and Sharon Simmons. Backward design: An integrated approach to a systems curriculum. In *Proceedings of the 46th ACM Technical Symposium on Computer Science Education*, SIGCSE '15, page 30–35, New York, NY, USA, 2015. Association for Computing Machinery.

[342] Michael S. Kirkpatrick and Samantha Prins. Using the readiness assurance process and metacognition in an operating systems course. In *Proceedings of the 2015 ACM Conference on Innovation and Technology in Computer Science Education*, ITiCSE '15, page 183–188, New York, NY, USA, 2015. Association for Computing Machinery.

[343] Anastasia Kitsantas and Barry J Zimmerman. Comparing self-regulatory processes among novice, non-expert, and expert volleyball players: A microanalytic study. *Journal of applied sport psychology*, 14(2):91–105, 2002.

[344] Erica Kleinman, Sabbir Ahmad, Zhaoqing Teng, Andy Bryant, Truong-Huy D. Nguyen, Casper Harteveld, and Magy Seif El-Nasr. "and then they died": Using action sequences for data driven, context aware gameplay analysis. In *International Conference on the Foundations of Digital Games*, FDG '20, New York, NY, USA, 2020. Association for Computing Machinery.

[345] Erica Kleinman, Christian Gayle, and Magy Seif El-Nasr. "because i'm bad at the game!" a microanalytical study of self regulated learning in league of legends. *Frontiers in Psychology*, 2021.

[346] Erica Kleinman, Reza Habibi, Yichen Yao, Christian Gayle, and Magy Seif El-Nasr. " a time and phase for everything"-towards a self regulated learning perspective on computational support for esports. *Proceedings of the ACM on Human-Computer Interaction*, 6(CHI PLAY):1–27, 2022.

[347] Erica Kleinman, Nikitha Preetham, Zhaoqing Teng, Andy Bryant, and Magy Seif El-Nasr. " what happened here!?" a taxonomy for user interaction with spatio-temporal game data visualization. *Proceedings of the ACM on Human-Computer Interaction*, 5(CHI PLAY):1–27, 2021.

[348] Erica Kleinman, Murtuza Shergadwala, Zhaoqing Teng, Jennifer Villareale, Andy Bryant, Jichen Zhu, and Magy Seif El-Nasr. Analyzing students' problem-solving sequences: A human-in-the-loop approach. *Journal of learning analytics*, 9(2):138–160, 2022.

[349] Erica Kleinman, Murtuza N Shergadwala, and Magy Seif El-Nasr. Kills, deaths, and (computational) assists: Identifying opportunities for computational support in esport learning. In *CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2022.

[350] Erica Kleinman, Jennifer Villareale, Murtuza Shergadwala, Zhaoqing Teng, Andy Bryant, Jichen Zhu, and Magy Seif El-Nasr. Towards an understanding of how

players make meaning from post-play process visualizations. In *International Conference on Entertainment Computing*, pages 47–58. Springer, 2022.

[351] Erica Seif Kleinman and Magy El-Nasr. Interactive visualization for strategy acquisition in esports spectatorship. In *Seegames Workshop, CHI 2020*, 2020.

[352] Bart P Knijnenburg, Svetlin Bostandjiev, John O'Donovan, and Alfred Kobsa. Inspectability and control in social recommenders. In *Proceedings of the sixth ACM conference on Recommender systems*, pages 43–50, 2012.

[353] Edwin M. Knorr. Reforming a database course to address cognitive load by using worked examples. *Proceedings of the 24th Western Canadian Conference on Computing Education, WCCCE 2019*, 2019.

[354] Daniel Knox and Sally Fincher. Where students go for knowledge and what they find there. *ICER 2013 - Proceedings of the 2013 ACM Conference on International Computing Education Research*, pages 35–40, 2013.

[355] Amy J. Ko, Thomas D. LaToza, Stephen Hull, Ellen A. Ko, William Kwok, Jane Quichocho, Harshitha Akkaraju, and Rishin Pandit. Teaching Explicit Programming Strategies to Adolescents. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education - SIGCSE '19*, pages 469–475, New York, New York, USA, Feb 2019. ACM Press.

[356] Athanasios Vasileios Kokkinakis, Simon Demediuk, Isabelle Nölle, Oluseyi Olarewaju, Sagarika Patra, Justus Robertson, Peter York, Alan Pedrassoli Pedrassoli Chitayat, Alistair Coates, Daniel Slawson, et al. Dax: Data-driven audience experiences in esports. In *ACM International Conference on Interactive Media Experiences*, pages 94–105, 2020.

[357] Alice Y. Kolb, David A. Kolb, Alice Y. Kolb, and David A. Kolb. The learning way: Meta-cognitive aspects of experiential learning. *Simulation & Gaming*, 40(3):297–327, 2009.

[358] David Kolb. *Experiential Learning: Experience as the Source of Learning and Development.* Prentice Hall, Englewoods Cliffs, NJ, 1983.

[359] David A Kolb, Richard E Boyatzis, Charalampos Mainemelis, et al. Experiential learning theory: Previous research and new directions. *Perspectives on thinking, learning, and cognitive styles*, 1(8):227–247, 2001.

[360] Denis Koposov, Maria Semenova, Andrey Somov, Andrey Lange, Anton Stepanov, and Evgeny Burnaev. Analysis of the reaction time of esports players through the gaze tracking and personality trait. In *2020 IEEE 29th International Symposium on Industrial Electronics (ISIE)*, pages 1560–1565. IEEE, 2020.

[361] Yubo Kou and Xinning Gui. Entangled with numbers: Quantified self and others in a team-based online game. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW):1–25, 2018.

[362] Yubo Kou and Xinning Gui. Emotion regulation in esports gaming: A qualitative study of league of legends. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW2):1–25, 2020.

[363] Matthias Kramer, David A. Tobinski, and Torsten Brinda. On the way to a test instrument for object-oriented programming competencies. In *Proceedings of the 16th Koli Calling International Conference on Computing Education Research - Koli Calling '16*, pages 145–149, New York, New York, USA, 2016. ACM Press.

[364] Ulrike-Marie Krause and Robin Stark. Reflection in example-and problem-based learning: Effects of reflection prompts, feedback and cooperative learning. *Evaluation & research in education*, 23(4):255–272, 2010.

[365] Sophia Krause-Levy, Leo Porter, Beth Simon, and Christine Alvarado. Investigating the impact of employing multiple interventions in a cs1 course. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*, SIGCSE

'20, page 1082–1088, New York, NY, USA, 2020. Association for Computing Machinery.

[366] Simone Kriglstein, Günter Wallner, and Margit Pohl. A user study of different gameplay visualizations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 361–370, 2014.

[367] Yen-Ting Kuan, Yu-Shuen Wang, and Jung-Hong Chuang. Visualizing real-time strategy games: The example of starcraft ii. In *2017 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pages 71–80. IEEE, 2017.

[368] S Kuhnt and A Krampe. An application of algebraic statistics to bowker's test for symmetry. In *14th INFORMS Applied Probability Conference (APS 2007)*, page 108. Eurandom & TU/e, 2007.

[369] Josef Künsting, Joachim Wirth, and Fred Paas. The goal specificity effect on strategy use and instructional efficiency during computer-based scientific discovery learning. *Computers & Education*, 56(3):668–679, 2011.

[370] Jaakko Kurhila, Miikka Miettinen, Petri Nokelainen, Patrik Floréen, and Henry Tirri. Peer-to-Peer Learning with Open-Ended Writable Web. *Proceedings of the Annual SIGCSE Conference on Innovation and Technology in Computer Science Education (ITiSCE)*, 8:173–177, 2003.

[371] Stan Kurkovsky, Stephanie Ludi, and Linda Clark. Active learning with lego for software requirements. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education*, SIGCSE '19, page 218–224, New York, NY, USA, 2019. ACM.

[372] Stan Kurkovsky, Stephanie Ludi, and Linda Clark. Active Learning with LEGO for Software Requirements. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education - SIGCSE '19*, pages 218–224, New York, New York, USA, 2019. ACM Press.

[373] Einari Kurvinen, Rolf Lindén, Teemu Rajala, Erkki Kaila, Mikko-Jussi Laakso, and Tapio Salakoski. Computer-assisted learning in primary school mathematics using ville education tool. In *Proceedings of the 12th Koli Calling International Conference on Computing Education Research*, Koli Calling '12, page 39–46, New York, NY, USA, 2012. Association for Computing Machinery.

[374] Mila Kwiatkowska. Measuring the Difficulty of Test Items in Computing Science Education. In *Proceedings of the 21st Western Canadian Conference on Computing Education - WCCCE '16*, pages 1–6, New York, New York, USA, May 2016. ACM Press.

[375] J Richard Landis and Gary G Koch. The measurement of observer agreement for categorical data. *biometrics*, pages 159–174, 1977.

[376] Janice Langan-Fox, Jeromy Anglim, and John R Wilson. Mental models, team mental models, and performance: Process, development, and future directions. *Human Factors and Ergonomics in Manufacturing & Service Industries*, 14(4):331–352, 2004.

[377] Lasse Juel Larsen. The play of champions: Toward a theory of skill in esport. *Sport, ethics and philosophy*, pages 1–23, 2020.

[378] Jean Lave and Etienne Wenger. *Situated learning: Legitimate peripheral participation*. Cambridge university press, 1991.

[379] Check-Yee Law, John Grundy, Andrew Cain, Rajesh Vasa, and Alex Cummaudo. User perceptions of using an open learner model visualisation tool for facilitating self-regulated learning. In *Proceedings of the Nineteenth Australasian Computing Education Conference*, pages 55–64, 2017.

[380] Fotis Lazarinis and Symeon Retalis. Analyze me: open learner model in an adaptive web testing system. *International Journal of Artificial Intelligence in Education*, 17(3):255–271, 2007.

[381] Richard S Lazarus and Susan Folkman. *Stress, appraisal, and coping.* Springer publishing company, 1984.

[382] Je Seok Lee, Minerva Wu, Diana Lee, Lee Fleming, Lindsay Ruben, Tom Turner, Kevin Brown, and Constance Steinkuehler. Designing an interest-based integrated curriculum around esports. *International Journal of Designs for Learning*, 11(3):78–95, 2020.

[383] Priscilla Lee and Soohyun Nam Liao. Targeting metacognition by incorporating student-reported confidence estimates on self-assessment quizzes. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education*, SIGCSE '21, page 431–437, New York, NY, USA, 2021. Association for Computing Machinery.

[384] Sukwon Lee, Sung-Hee Kim, Ya-Hsin Hung, Heidi Lam, Youn-ah Kang, and Ji Soo Yi. How do people make sense of unfamiliar visualizations?: A grounded model of novice's information visualization sensemaking. *IEEE transactions on visualization and computer graphics*, 22(1):499–508, 2015.

[385] Tsang-Hsiung Lee, Pei-Di Shen, and Chia-Wen Tsai. Enhance low-achieving students' learning involvement in taiwan's higher education: an approach via e-learning with problem-based learning and self-regulated learning. *Teaching in Higher Education*, 15(5):553–565, 2010.

[386] Äli Leijen, Kai Valtna, Djuddah AJ Leijen, and Margus Pedaste. How to determine the quality of students' reflections? *Studies in Higher Education*, 37(2):203–217, 2012.

[387] Noel LeJeune. Critical components for successful collaborative learning in cs1. *Journal of Computing Sciences in Colleges*, 19(1):275–285, 2003.

[388] Leo Leppänen, Juho Leinonen, and Arto Hellas. Pauses and spacing in learning to program. In *Proceedings of the 16th Koli Calling International Conference on*

Computing Education Research, Koli Calling '16, page 41–50, New York, NY, USA, 2016. Association for Computing Machinery.

[389] Leo Leppänen, Juho Leinonen, and Arto Hellas. Pauses and spacing in learning to program. In *Proceedings of the 16th Koli Calling International Conference on Computing Education Research - Koli Calling '16*, pages 41–50, New York, New York, USA, Nov 2016. ACM Press.

[390] Eric L. Lesser and John Storck. Communities of practice and organizational performance. *IBM systems journal*, 40(4):831–841, 2001.

[391] Gary. Lewandowski and ACM. Special Interest Group on Computer Science Education. *SIGCSE '10 : Proceedings of the 41st ACM Technical Symposium on Computer Science Education, Milwaukee, Wisconsin, USA, March 10-13, 2010.* ACM, 2010.

[392] Gary. Lewandowski and ACM. Special Interest Group on Computer Science Education. *SIGCSE '10 : Proceedings of the 41st ACM Technical Symposium on Computer Science Education, Milwaukee, Wisconsin, USA, March 10-13, 2010.* ACM, 2010.

[393] Soohyun Nam Liao, Sander Valstar, Kevin Thai, Christine Alvarado, Daniel Zingaro, William G. Griswold, and Leo Porter. Behaviors of Higher and Lower Performing Students in CS1. In *Proceedings of the 2019 ACM Conference on Innovation and Technology in Computer Science Education - ITiCSE '19*, pages 196–202, New York, New York, USA, 2019. ACM Press.

[394] Beng-Chong Lim and Katherine J Klein. Team mental models and team performance: A field study of the effects of team mental model similarity and accuracy. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, 27(4):403–418, 2006.

[395] Viktoriya Lipovaya, Yuri Lima, Pedro Grillo, Carlos Eduardo Barbosa, Jano Moreira de Souza, and Francisco José de Castro Moura Duarte. Coordination, communication, and competition in esports: A comparative analysis of teams in two action games. In *Proceedings of 16th European Conference on Computer-Supported Cooperative Work-Exploratory Papers*. European Society for Socially Embedded Technologies (EUSSET), 2018.

[396] Alex Lishinski, Aman Yadav, Jon Good, and Richard Enbody. Learning to program: Gender differences and interactive effects of students' motivation, goals, and self-efficacy on performance. In *Proceedings of the 2016 ACM Conference on International Computing Education Research*, ICER '16, page 211–220, New York, NY, USA, 2016. Association for Computing Machinery.

[397] Hsu Kuan Jonathan LIU. Correlation research on the application of e-learning to students' self-regulated learning ability, motivational beliefs, and academic performance. *EURASIA Journal of Mathematics, Science and Technology Education*, 12(4):1091–1100, 2016.

[398] Zhicheng Liu and John Stasko. Mental models, visual reasoning and interaction in information visualization: A top-down perspective. *IEEE transactions on visualization and computer graphics*, 16(6):999–1008, 2010.

[399] Richard Lobb and Jenny Harlow. Coderunner: A tool for assessing computer programming skills. *ACM Inroads*, 7(1):47–51, feb 2016.

[400] Dastyni Loksa and Amy J. Ko. The role of self-regulation in programming problem solving process and success. In *Proceedings of the 2016 ACM Conference on International Computing Education Research*, ICER '16, page 83–91, New York, NY, USA, 2016. Association for Computing Machinery.

[401] Dastyni Loksa, Amy J. Ko, Will Jernigan, Alannah Oleson, Christopher J. Mendez, and Margaret M. Burnett. Programming, problem solving, and self-

awareness: Effects of explicit guidance. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, page 1449–1461, New York, NY, USA, 2016. Association for Computing Machinery.

[402] Dastyni Loksa, Lauren Margulieux, Brett A. Becker, Michelle Craig, Paul Denny, Raymond Pettit, and James Prather. Metacognition and self-regulation in programming education: Theories and exemplars of use. *ACM Trans. Comput. Educ.*, Dec 2021. Just Accepted.

[403] Dastyni Loksa, Benjamin Xie, Harrison Kwik, and Amy J. Ko. Investigating novices' in situ reflections on their programming process. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*, SIGCSE '20, page 149–155, New York, NY, USA, 2020. Association for Computing Machinery.

[404] Yanjin Long and Vincent Aleven. Enhancing learning outcomes through self-regulated learning support with an open learner model. *User Modeling and User-Adapted Interaction*, 27(1):55–88, 2017.

[405] Marsha Lovett. Make exams worth more than the grade. In Danielle LaVaque-Manty Deborah Meizlish Matthew Kaplan, Naomi Silver, editor, *Using Reflection and Metacognition to Improve Student Learning: Across the Disciplines, Across the Academy*, chapter 2, pages 18–52. Stylus, Virginia, 2013.

[406] Andrew Luxton-Reilly, Paul Denny, Beryl Plimmer, and Robert Sheehan. Activities, affordances and attitude. In *Proceedings of the 17th ACM annual conference on Innovation and technology in computer science education - ITiCSE '12*, page 4, New York, New York, USA, 2012. ACM Press.

[407] Andrew Luxton-Reilly, Judy Sheard, Claudia Szabo, Simon, Ibrahim Albluwi, Brett A. Becker, Michail Giannakos, Amruth N. Kumar, Linda Ott, James Paterson, and Michael James Scott. Introductory programming: a systematic literature review. In *Proceedings Companion of the 23rd Annual ACM Conference on Inno-*

vation and Technology in Computer Science Education - ITiCSE 2018 Companion, pages 55–106, New York, New York, USA, Jul 2018. ACM Press.

[408] Andrew Luxton-Reilly, Simon, Ibrahim Albluwi, Brett A. Becker, Michail Giannakos, Amruth N. Kumar, Linda Ott, James Paterson, Michael James Scott, Judy Sheard, and et al. Introductory programming: A systematic literature review. In *Proceedings Companion of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education*, ITiCSE 2018 Companion, page 55–106, New York, NY, USA, 2018. ACM.

[409] Brooke N Macnamara, David Moreau, and David Z Hambrick. The relationship between deliberate practice and performance in sports: A meta-analysis. *Perspectives on Psychological Science*, 11(3):333–350, 2016.

[410] Stephen MacNeil, Celine Latulipe, Bruce Long, and Aman Yadav. Exploring Lightweight Teams in a Distributed Learning Environment. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education - SIGCSE '16*, pages 193–198, New York, New York, USA, 2016. ACM Press.

[411] Stephen MacNeil, Celine Latulipe, and Aman Yadav. Learning in distributed low-stakes teams. *ICER 2015 - Proceedings of the 2015 ACM Conference on International Computing Education Research*, pages 227–236, 2015.

[412] Daniel Madden, Yuxuan Liu, Haowei Yu, Mustafa Feyyaz Sonbudak, Giovanni M Troiano, and Casper Harteveld. "why are you playing games? you are a girl!": Exploring gender biases in esports. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–15, 2021.

[413] Carlo Magno. Assessing academic self-regulated learning among filipino college students: The factor structure and item fit. *The international Journal of Educational and psychological assessment*, 5, 2010.

[414] Dara Mahdid, Piotr Szybek, and Führer. A study on variation technique in courses on scientific computing. *Science Journal of Education*, 3(3):60–67, 2015.

[415] Pekka Makkonen. Does collaborative hypertext support better engagement in learning of the basics in informatics? *ACM SIGCSE Bulletin*, 29(3):130–132, Sep 1997.

[416] Jonna Malmberg, Sanna Järvelä, and Hanna Järvenoja. Capturing temporal and sequential patterns of self-, co-, and socially shared regulation in the context of collaborative learning. *Contemporary Educational Psychology*, 49:160–174, 2017.

[417] Lauri Malmi, Judy Sheard, Päivi Kinnunen, Simon, and Jane Sinclair. Computing education theories: What are they and how are they used? In *Proceedings of the 2019 ACM Conference on International Computing Education Research*, ICER '19, page 187–197, New York, NY, USA, 2019. ACM.

[418] Lauri Malmi, Judy Sheard, Päivi Kinnunen, and Jane Sinclair. Theories and models of emotions, attitudes, and self-efficacy in the context of programming education. In *Proceedings of the 2020 ACM Conference on International Computing Education Research*, pages 36–47, 2020.

[419] Lauri Malmi, Judy Sheard, Simon, Roman Bednarik, Juha Helminen, Päivi Kinnunen, Ari Korhonen, Niko Myller, Juha Sorva, and Ahmad Taherkhani. Theoretical underpinnings of computing education research: What is the evidence? In *Proceedings of the Tenth Annual Conference on International Computing Education Research*, ICER '14, page 27–34, New York, NY, USA, 2014. ACM.

[420] Katerina Mangaroska, Kshitij Sharma, Dragan Gasevic, and Michail Giannakos. Exploring students' cognitive and affective states during problem solving through multimodal data: Lessons learned from a programming activity. *Journal of Computer Assisted Learning*, 38(1):40–59, 2022.

[421] Murali Mani and Quamrul Mazumder. Incorporating metacognition into learning. In *Proceedings of the 44th ACM Technical Symposium on Computer Science Education*, SIGCSE '13, page 53–58, New York, NY, USA, 2013. Association for Computing Machinery.

[422] Linda Mannila, Valentina Dagiene, Barbara Demo, Natasa Grgurina, Claudio Mirolo, Lennart Rolandsson, and Amber Settle. Computational thinking in K-9 education. *ITiCSE-WGR 2014 - Working Group Reports of the 2014 Innovation and Technology in Computer Science Education Conference*, pages 1–29, 2014.

[423] Linda Mannila, Lars-Åke Nordén, and Arnold Pears. Digital Competence, Teacher Self-Efficacy and Training Needs. In *Proceedings of the 2018 ACM Conference on International Computing Education Research - ICER '18*, pages 78–85, New York, New York, USA, 2018. ACM Press.

[424] Linda Mannila, Lars-Åke Nordén, and Arnold Pears. Digital competence, teacher self-efficacy and training needs. In *Proceedings of the 2018 ACM Conference on International Computing Education Research*, ICER '18, page 78–85, New York, NY, USA, 2018. ACM.

[425] Michelle A. Marks, John E. Mathieu, and Stephen J. Zaccaro. A temporally based framework and taxonomy of team processes. *The Academy of Management Review*, 26(3):356–376, 2001.

[426] C Dianne Martin. The case for integrating ethical and social impact into the computer science curriculum. In *The supplemental Proceedings of the conference on Integrating technology into computer science education: working group reports and supplemental proceedings - ITiCSE-WGR '97*, pages 114–120, New York, New York, USA, 1997. ACM Press.

[427] Joshua Martin, Stephen H. Edwards, and Clfford A. Shaffer. The effects of procrastination interventions on programming project success. In *Proceedings of the*

Eleventh Annual International Conference on International Computing Education Research, ICER '15, page 3–11, New York, NY, USA, 2015. Association for Computing Machinery.

[428] Nikolas Martinos. Functional and non-functional requirements for a performance dashboard in cs: Go. B.S. thesis, University of Twente, 2021.

[429] Ference Marton. *Describing and Improving Learning*, pages 53–82. Springer US, Boston, MA, 1988.

[430] Samiha Marwan, Anay Dombe, and Thomas W. Price. Unproductive help-seeking in programming: What it is and how to address it. In *Proceedings of the 2020 ACM Conference on Innovation and Technology in Computer Science Education*, ITiCSE '20, page 54–60, New York, NY, USA, 2020. Association for Computing Machinery.

[431] Susana Masapanta-Carrión and J. Ángel Velázquez-Iturbide. Evaluating instructors' classification of programming exercises using the revised Bloom's taxonomy. In *Annual Conference on Innovation and Technology in Computer Science Education, ITiCSE*, pages 541–547. ACM, Jul 2019.

[432] Eva Mayr, Günther Schreder, Michael Smuc, and Florian Windhager. Looking at the representations in our mind: Measuring mental models of information visualizations. In *Proceedings of the Sixth Workshop on Beyond Time and Errors on Novel Evaluation Methods for Visualization*, pages 96–103, 2016.

[433] Lindsay McCardle. *Self-regulated learning in and across sport and academic domains*. PhD thesis, 2015.

[434] Lindsay McCardle. Similarities and differences in self-regulated learning processes in sport and academics: A case study. *Journal for the Study of Sports and Athletes in Education*, 9(3):190–213, 2015.

[435] Paul J McCarthy, Marc V Jones, Chris G Harwood, and Laura Davenport. Using goal setting to enhance positive affect among junior multievent athletes. *Journal of Clinical Sport Psychology*, 4(1):53–68, 2010.

[436] Mary McCaslin and Ruby Inez Vega. Peer co-regulation of learning, emotion, and coping in small-group learning. In *Constructing Educational Achievement*, pages 118–135. Routledge, 2014.

[437] Renee McCauley, Tracy Camp, Paul Tymon, J. D. Dougherty, Kris Nagel, and ACM. Special Interest Group on Computer Science Education. *SIGCSE'13 : Proceedings of the 44th ACM Technical Symposium on Computer Science Education : March 6-9, 2013, Denver, Colorado, USA.*

[438] Michael McCracken and Robert Waters. Why? When an otherwise successful intervention fails. *SIGCSE Bulletin (ACM, Special Interest Group on Computer Science Education)*, 31(3):9–12, 1999.

[439] Roger McDermott, Mats Daniels, and Åsa Cajander. Perseverance Measures and Attainment in First Year Computing Science Students. In *Proceedings of the 2015 ACM Conference on Innovation and Technology in Computer Science Education - ITiCSE '15*, volume 2015-June, pages 302–307, New York, New York, USA, Jun 2015. ACM Press.

[440] Roger McDermott, Iain Pirie, Åsa Cajander, Mats Daniels, and Cary Laxer. Investigation into the personal epistemology of computer science students. In *Proceedings of the 18th ACM conference on Innovation and technology in computer science education - ITiCSE '13*, page 231, New York, New York, USA, 2013. ACM Press.

[441] Roger McDermott, Iain Pirie, Åsa Cajander, Mats Daniels, and Cary Laxer. Investigation into the personal epistemology of computer science students. In *Proceedings of the 18th ACM Conference on Innovation and Technology in Computer Science Education*, ITiCSE '13, page 231–236, New York, NY, USA, 2013. ACM.

[442] Roger McDermott, Mark Zarb, Mats Daniels, Åsa Cajander, and Tony Clear. Motivation, Optimal Experience and Flow in First Year Computing Science. In *Proceedings of the 2016 ACM Conference on Innovation and Technology in Computer Science Education - ITiCSE '16*, volume 11-13-July, pages 206–211, New York, New York, USA, Jul 2016. ACM Press.

[443] Monica M. McGill, Tom McKlin, Adrienne Decker, and Kathy Haynie. A gap analysis of noncognitive constructs in evaluation instruments designed for computing education. *SIGCSE 2019 - Proceedings of the 50th ACM Technical Symposium on Computer Science Education*, pages 706–712, 2019.

[444] Dawn McKinney and Leo F Denton. Houston, we have a problem: There's a leak in the CS1 affective oxygen tank. In *Proceedings of the 35th SIGCSE technical symposium on Computer science education - SIGCSE '04*, page 236, New York, New York, USA, 2004. ACM Press.

[445] Tom McKlin, Brian Magerko, Taneisha Lee, Dana Wanzer, Doug Edwards, and Jason Freeman. Authenticity and Personal Creativity. In *Proceedings of the 49th ACM Technical Symposium on Computer Science Education - SIGCSE '18*, volume 2018-Janua, pages 987–992, New York, New York, USA, 2018. ACM Press.

[446] Lilla Dale McManis and Susan B Gunnewig. Finding the education in educational technology with early learners. *Young children*, 67(3):14–24, 2012.

[447] Ben Medler. Player dossiers: Analyzing gameplay data as a reward. *Game Studies*, 11(1), 2011.

[448] Ben Medler, Michael John, and Jeff Lane. Data cracker: developing a visual game analytic tool for analyzing online gameplay. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 2365–2374, 2011.

[449] Ben Medler and Brian Magerko. Analytics of play: Using information visualization

and gameplay practices for visualizing video game data. *Parsons Journal for Information Mapping*, 3(1):1–12, 2011.

[450] David A Medler. A brief history of connectionism. *Neural computing surveys*, 1:18–72, 1998.

[451] M. Dee Medley, Rebecca H. Rutherfoord, G. Ernest Anderson, R. Waldo Roth, and Stuart A. Varden. Ethical issues related to internet development and research. In *Working Group reports of the 3rd annual SIGCSE/SIGCUE ITiCSE conference on Integrating technology into computer science education - ITiCSE-WGR '98*, pages 57–72, New York, New York, USA, 1998. ACM Press.

[452] Orni Meerbaum-Salant, Michal Armoni, and Mordechai (Moti) Ben-Ari. Learning computer science concepts with scratch. In *Proceedings of the Sixth international workshop on Computing education research - ICER '10*, page 69, New York, New York, USA, 2010. ACM Press.

[453] Elisa D Mekler, Ioanna Iacovides, and Julia Ayumi Bopp. " a game that makes you question..." exploring the role of reflection for the player experience. In *Proceedings of the 2018 annual symposium on computer-human interaction in play*, pages 315–327, 2018.

[454] Muhsin Menekse, Saira Anwar, and Zeynep Gonca Akdemir. How do different reflection prompts affect engineering students' academic performance and engagement? *The Journal of Experimental Education*, 90(2):261–279, 2022.

[455] Jack Mezirow. A critical theory of adult learning and education. *Adult education*, 32(1):3–24, 1981.

[456] Lan Min and Hew Khe Foon. Self-regulated learning process in moocs: Examining the indicators of behavioral, emotional, and cognitive engagement. In *Proceedings of the 2019 4th International Conference on Distance Education and Learning*, pages 99–105, 2019.

[457] A MITROVIC. Evaluation of constraint-based tutor for database. *International Journal of AIED*, 10(3):238–256, 1999.

[458] Antonija Mitrovic. Fifteen years of constraint-based tutors: what we have achieved and where we are going. *User modeling and user-adapted interaction*, 22(1-2):39–72, 2012.

[459] mobalytics. Mobalytics. https://app.mobalytics.gg/, 2021. Online; accessed 25 January 2021.

[460] Susan Mohammed and Brad C Dumville. Team mental models in a team knowledge framework: Expanding theory and measurement across disciplinary boundaries. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, 22(2):89–106, 2001.

[461] Trevor T Moores, Jerry Cha-Jan Chang, and Deborah K Smith. Clarifying the role of self-efficacy and metacognition as predictors of performance: Construct development and test. *ACM SIGMIS Database: the DATABASE for Advances in Information Systems*, 37(2-3):125–132, 2006.

[462] Andrés Moreno, Erkki Sutinen, and Mike Joy. Defining and evaluating conflictive animations for programming education: The case of Jeliot ConAn. In *SIGCSE 2014 - Proceedings of the 45th ACM Technical Symposium on Computer Science Education*, pages 629–634. ACM, 2014.

[463] Michael Morgan, Matthew Butler, Jane Sinclair, Christabel Gonsalvez, and Neena Thota. Contrasting CS student and academic perspectives and experiences of student engagement. In *Proceedings Companion of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education - ITiCSE 2018 Companion*, pages 1–35, New York, New York, USA, 2018. ACM Press.

[464] Michael Morgan, Matthew Butler, Neena Thota, and Jane Sinclair. How CS academics view student engagement. In *Proceedings of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education - ITiCSE 2018*, pages 284–289, New York, New York, USA, Jul 2018. ACM Press.

[465] Michael Morgan, Jane Sinclair, Matthew Butler, Neena Thota, Janet Fraser, Gerry Cross, and Jana Jackova. Understanding International Benchmarks on Student Engagement. In *Proceedings of the 2017 ITiCSE Conference on Working Group Reports - ITiCSE-WGR '17*, volume 2018-Janua, pages 1–24, New York, New York, USA, 2017. ACM Press.

[466] Adon Christian Michael Moskal and Rob Wass. Interpersonal process recall: A novel approach to illuminating students' software development processes. *Computer Science Education*, 29(1):5–22, 2019.

[467] Dinara Moura, Magy Seif el Nasr, and Christopher D Shaw. Visualizing and understanding players' behavior in video games: discovering patterns and supporting aggregation and comparison. In *Proceedings of the 2011 ACM SIGGRAPH symposium on video games*, pages 11–15. ACM, 2011.

[468] Erica Mourão, João Felipe Pimentel, Leonardo Murta, Marcos Kalinowski, Emilia Mendes, and Claes Wohlin. On the performance of hybrid search strategies for systematic literature reviews in software engineering. *Information and Software Technology*, 123:106294, 2020.

[469] Laurie Murphy, Timothy Fossum, Susan Haller, Kate Sanders, Renée McCauley, Briana B. Morrison, Carol Zander, Suzanne Westbrook, Brad Richards, and Ruth E. Anderson. A multi-institutional investigation of computer science seniors' knowledge of programming concepts. *Proceedings of the Thirty-Sixth SIGCSE Technical Symposium on Computer Science Education, SIGCSE 2005*, pages 510–514, 2005.

[470] Laurie Murphy and Josh Tenenberg. Do computer science students know what they know? a calibration study of data structure knowledge. In *Proceedings of the 10th Annual SIGCSE Conference on Innovation and Technology in Computer Science Education*, ITiCSE '05, page 148–152, New York, NY, USA, 2005. Association for Computing Machinery.

[471] Nicholas P Murray. The differential effect of team cohesion and leadership behavior in high school sports. *Individual Differences Research*, 4(4), 2006.

[472] Ilya Musabirov, Paul Okopny, and Denis Bulygin. Analyzing chat logs in online games for tutorial improvement. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*, pages 661–666, 2015.

[473] Geoff Musick, Rui Zhang, Nathan J McNeese, Guo Freeman, and Anurata Prabha Hridi. Leveling up teamwork in esports: Understanding team cognition in a dynamic virtual environment. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW1):1–30, 2021.

[474] Ratna Narayan, Cynthia Rodriguez, Juan Araujo, Ali Shaqlaih, and Glenda Moss. Constructivism—constructivist learning theory. 2013.

[475] Farzaneh Nasirian, Mohsen Ahmadian, and One-Ki Daniel Lee. Ai-based voice assistant systems: Evaluating from the interaction and trust perspectives. 2017.

[476] Greg L. Nelson and Amy J. Ko. On use of theory in computing education research. In *Proceedings of the 2018 ACM Conference on International Computing Education Research*, ICER '18, page 31–39, New York, NY, USA, 2018. ACM.

[477] Liat Netzer, Gerben A Van Kleef, and Maya Tamir. Interpersonal instrumental emotion regulation. *Journal of experimental social psychology*, 58:124–135, 2015.

[478] Truong-Huy D Nguyen, Magy Seif El-Nasr, and Alessandro Canossa. Glyph: Visualization tool for understanding problem solving strategies in puzzle games. In *FDG*, 2015.

[479] Rune Kristian Lundedal Nielsen and Thorkild Hanghøj. Esports skills are people skills. In *Proceedings of the 13th European Conference on Game-Based Learning*, pages 535–542, 2019.

[480] John L Nietfeld. The role of self-regulated learning in digital games. In *Handbook of self-regulation of learning and performance*, pages 271–284. Routledge, 2017.

[481] John L Nietfeld, Lucy R Shores, and Kristin F Hoffmann. Self-regulation and gender within a game-based learning environment. *Journal of Educational Psychology*, 106(4):961, 2014.

[482] The Language Archive Nijmegen: Max Planck Institute for Psycholinguistics. Elan (version 6.0) [computer software]. Retrieved from https://archive.mpi.nl/tla/elan, 2020.

[483] Susan Nolen-Hoeksema. Emotion regulation and psychopathology: The role of gender. *Annual review of clinical psychology*, 8:161–187, 2012.

[484] Henrik Nygren, Juho Leinonen, and Arto Hellas. Non-restricted Access to Model Solutions. In *Proceedings of the 2019 ACM Conference on Innovation and Technology in Computer Science Education - ITiCSE '19*, pages 44–50, New York, New York, USA, 2019. ACM Press.

[485] Kevin O'Kelley. New employees and safety culture: A social cognitive theory perspective. *Professional Safety*, 64(02):37–40, 2019.

[486] Aleksi Oksala. The importance of communication in esports: an ethnographic interview with a finnish counter-strike: Global offensive team. 2022.

[487] Martin Oliver and Diane Carr. Learning in virtual worlds: Using communities of practice to explain how people learn from play. *British journal of educational technology*, 40(3):444–457, 2009.

[488] Jennifer K. Olsen and Armando Fox. Usage of hints on coding-based summative assessments. In *SIGCSE 2019 - Proceedings of the 50th ACM Technical Symposium on Computer Science Education*, pages 839–844. ACM, Inc, Feb 2019.

[489] op.gg. op.gg. https://na.op.gg/, 2021. Online; accessed 25 January 2021.

[490] R. Brook Osborne, Antony J. Thomas, and Jeffrey R.N. Forbes. Teaching with robots: A service-learning approach to mentor training. *SIGCSE'10 - Proceedings of the 41st ACM Technical Symposium on Computer Science Education*, pages 172–176, 2010.

[491] Claudia Ott, Anthony Robins, Patricia Haden, and Kerry Shephard. Illustrating performance indicators and course characteristics to support students' self-regulated learning in cs1. *Computer Science Education*, 25(2):174–198, 2015.

[492] Overwatch Tracker. OVERWATCH TRACKER. https://https://overwatchtracker.com/, 2021. Online; accessed 25 January 2021.

[493] Harold F O'Neil, Gregory KWK Chung, Deirdre Kerr, Terry P Vendlinski, Rebecca E Buschang, and Richard E Mayer. Adding self-explanation prompts to an educational computer game. *Computers in Human Behavior*, 30:23–28, 2014.

[494] F. Pajares. Current directions in self-efficacy research. In M. Maehr and P. R. Pintrich, editors, *Advances in Motivation and Achievement*, pages 1–49. JAI Press, Greenwich, CT, 1997.

[495] Annemarie Palincsar and Ann Brown. Reciprocal teaching of comprehension-monitoring activities. *Technical Report No. 269.*, 1983.

[496] Sonia Pamplona, Nelson Medinilla, and Pamela Flores. Exploring misconceptions of operating systems in an online course. In *Proceedings of the 13th Koli Calling International Conference on Computing Education Research - Koli Calling '13*, pages 77–86, New York, New York, USA, 2013. ACM Press.

[497] Ernesto Panadero. A review of self-regulated learning: Six models and four directions for research. *Frontiers in psychology*, 8:422, 2017.

[498] Zacharoula K Papamitsiou and Anastasios A Economides. Learning analytics and educational data mining in practice: A systematic literature review of empirical evidence. *Educational Technology & Society*, 17(4):49–64, 2014.

[499] Jennifer Parham, Leo Gugerty, and D. E. Stevenson. Empirical evidence for the existence and uses of metacognition in computer science problem solving. In *Proceedings of the 41st ACM Technical Symposium on Computer Science Education*, SIGCSE '10, page 416–420, New York, NY, USA, 2010. Association for Computing Machinery.

[500] S. G. Paris and P. Winograd. How metacognition can promote academic learning and instruction. In B. F. Jones and L. Idol, editors, *Dimensions of Thinking and Cognitive Instruction*, pages 15–51. Lawrence Erlbaum Associates, Inc., 1990.

[501] Yeonjeong Park and I-H Jo. Development of the learning analytics dashboard to support students' learning performance. *Journal of Universal Computer Science*, 21(1):110, 2015.

[502] Miranda C. Parker, Kantwon Rogers, Barbara J. Ericson, and Mark Guzdial. Students and teachers use an online ap cs principles ebook differently: Teacher behavior consistent with expert learners. In *Proceedings of the 2017 ACM Conference on International Computing Education Research*, ICER '17, page 101–109, New York, NY, USA, 2017. Association for Computing Machinery.

[503] Petr Parshakov, Dennis Coates, and Marina Zavertiaeva. Is diversity good or bad? evidence from esports teams analysis. *Applied Economics*, 50(47):5064–5075, 2018.

[504] Nathan Partlan, Abdelrahman Madkour, Chaima Jemmali, Josh Aaron Miller, Christoffer Holmgård, and Magy Seif El-Nasr. Player imitation for build actions

in a real-time strategy game. In *AIIDE workshop on Artificial Intelligence for Strategy Games*, 2019.

[505] Avi Parush, Chelsea Kramer, Tara Foster-Hunt, Kathryn Momtahan, Aren Hunter, and Benjamin Sohmer. Communication and team situation awareness in the or: Implications for augmentative information display. *Journal of biomedical informatics*, 44(3):477–485, 2011.

[506] Robert E Patterson, Leslie M Blaha, Georges G Grinstein, Kristen K Liggett, David E Kaveney, Kathleen C Sheldon, Paul R Havig, and Jason A Moore. A human cognition framework for information visualization. *Computers & Graphics*, 42:42–58, 2014.

[507] Allison W Pearson, Michael D Ensley, and Allen C Amason. An assessment and refinement of jehn's intragroup conflict scale. *International Journal of Conflict Management*, 2002.

[508] Ismael Pedraza-Ramirez, Lisa Musculus, Markus Raab, and Sylvain Laborde. Setting the scientific stage for esports psychology: A systematic review. *International Review of Sport and Exercise Psychology*, 13(1):319–352, 2020.

[509] Sean Penney, Jonathan Dodge, Claudia Hilderbrand, Andrew Anderson, Logan Simpson, and Margaret Burnett. Toward foraging for understanding of starcraft agents: An empirical study. In *23rd International Conference on Intelligent User Interfaces*, pages 225–237, 2018.

[510] Markeya S. Peteranetz, Abraham E. Flanigan, Duane F. Shell, and Leen-Kiat Soh. Perceived Instrumentality and Career Aspirations in CS1 Courses. In *Proceedings of the 2016 ACM Conference on International Computing Education Research - ICER '16*, pages 13–21, New York, New York, USA, 2016. ACM Press.

[511] Markeya S. Peteranetz, Patrick M. Morrow, and Leen-Kiat Soh. Development and validation of the computational thinking concepts and skills test. In *Proceedings of*

the 51st ACM Technical Symposium on Computer Science Education, SIGCSE '20, page 926–932, New York, NY, USA, 2020. Association for Computing Machinery.

[512] Markeya S. Peteranetz, Shiyuan Wang, Duane F. Shell, Abraham E. Flanigan, and Leen-Kiat Soh. Examining the impact of computational creativity exercises on college computer science students' learning, achievement, self-efficacy, and creativity. In *Proceedings of the 49th ACM Technical Symposium on Computer Science Education*, SIGCSE '18, page 155–160, New York, NY, USA, 2018. Association for Computing Machinery.

[513] Andrew Petersen, Michelle Craig, Jennifer Campbell, and Anya Tafliovich. Revisiting why students drop CS1. In *Proceedings of the 16th Koli Calling International Conference on Computing Education Research - Koli Calling '16*, pages 71–80, New York, New York, USA, 2016. ACM Press.

[514] Carl Adam Petri and Wolfgang Reisig. Petri net. *Scholarpedia*, 3(4):6477, 2008.

[515] Paul R. Pintrich. Chapter 14 - the role of goal orientation in self-regulated learning. In Monique Boekaerts, Paul R. Pintrich, and Moshe Zeidner, editors, *Handbook of Self-Regulation*, pages 451 – 502. Academic Press, San Diego, 2000.

[516] Paul R Pintrich and Elisabeth V De Groot. Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82(1):33–40, 1990.

[517] Paul R Pintrich et al. A manual for the use of the motivated strategies for learning questionnaire (mslq). 1991.

[518] Paul R. Pintrich, David A. F. Smith, Teresa Garcia, and Wilbert J. Mckeachie. Reliability and predictive validity of the motivated strategies for learning questionnaire (mslq). *Educational and Psychological Measurement*, 53(3):801–813, 1993.

[519] Peter Pirolli and Stuart Card. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of*

international conference on intelligence analysis, volume 5, pages 2–4. McLean, VA, USA, 2005.

[520] Matthew A Pluss, Andrew R Novak, KJ Bennett, Derek Panchuk, Aaron J Coutts, and Job Fransen. Perceptual-motor abilities underlying expertise in esports. *Journal of Expertise*, 3(2):133–143, 2020.

[521] Matthew A Pluss, Andrew R Novak, Kyle JM Bennett, Derek Panchuk, Aaron J Coutts, and Job Fransen. The relationship between the quantity of practice and in-game performance during practice with tournament performance in esports: An eight-week study.

[522] Matthew A Pluss, Andrew R Novak, Kyle JM Bennett, Derek Panchuk, Aaron J Coutts, and Job Fransen. The relationship between the quantity of practice and in-game performance during practice with tournament performance in esports: An eight-week study. *The Journal of Sport and Exercise Science*, 5(1):69–76, 2021.

[523] Joe Gibbs Politz, Daniel Patterson, Shriram Krishnamurthi, and Kathi Fisler. CaptainTeach: multi-stage, in-flow peer review for programming assignments. In *Proceedings of the 2014 conference on Innovation & technology in computer science education - ITiCSE '14*, pages 267–272, New York, New York, USA, 2014. ACM Press.

[524] Heather Pon-Barry, Audrey St. John, Becky Wai-Ling Packard, and Barbara Rotundo. A Flexible Curriculum for Promoting Inclusion through Peer Mentorship. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education - SIGCSE '19*, pages 1116–1122, New York, New York, USA, 2019. ACM Press.

[525] porofessor.gg. porofessor.gg. https://porofessor.gg/download, 2021. Online; accessed December 9th 2021.

[526] Ron Porter and Paul Calder. Patterns in learning to program: An experiment?

In *Proceedings of the Sixth Australasian Conference on Computing Education - Volume 30*, ACE '04, page 241–246, AUS, 2004. Australian Computer Society, Inc.

[527] Dylan R Poulus, Tristan J Coulter, Michael G Trotter, and Remco Polman. A qualitative analysis of the perceived determinants of success in elite esports athletes. *Journal of Sports Sciences*, 40(7):742–753, 2022.

[528] Zachary Pousman, John Stasko, and Michael Mateas. Casual information visualization: Depictions of data in everyday life. *IEEE transactions on visualization and computer graphics*, 13(6):1145–1152, 2007.

[529] Jean H Powell. The reflective practitioner in nursing. *Journal of Advanced Nursing*, 14(10):824–832, 1989.

[530] James Prather, Brett A. Becker, Michelle Craig, Paul Denny, Dastyni Loksa, and Lauren Margulieux. What do we think we think we are doing? metacognition and self-regulation in programming. In *Proceedings of the 2020 ACM Conference on International Computing Education Research*, ICER '20, page 2–13, New York, NY, USA, 2020. Association for Computing Machinery.

[531] James Prather, Lauren Margulieux, Jacqueline Whalley, Paul Denny, Brent N. Reeves, Brett A. Becker, Paramvir Singh, Garrett Powell, and Nigel Bosch. Getting by with help from my friends: Group study in introductory programming understood as socially shared regulation. In *Proceedings of the 2022 ACM Conference on International Computing Education Research - Volume 1*, ICER '22, page 164–176, New York, NY, USA, 2022. Association for Computing Machinery.

[532] James Prather, Lauren Margulieux, Jacqueline Whalley, Paul Denny, Brent N Reeves, Brett A Becker, Paramvir Singh, Garrett Powell, and Nigel Bosch. Getting by with help from my friends: Group study in introductory programming understood as socially shared regulation. In *Proceedings of the 2022 ACM Con-*

ference on International Computing Education Research-Volume 1, pages 164–176, 2022.

[533] James Prather, Raymond Pettit, Brett A. Becker, Paul Denny, Dastyni Loksa, Alani Peters, Zachary Albrecht, and Krista Masci. First things first: Providing metacognitive scaffolding for interpreting problem prompts. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education*, SIGCSE '19, page 531–537, New York, NY, USA, 2019. Association for Computing Machinery.

[534] James Prather, Raymond Pettit, Kayla McMurry, Alani Peters, John Homer, and Maxine Cohen. Metacognitive difficulties faced by novice programmers in automated assessment tools. In *Proceedings of the 2018 ACM Conference on International Computing Education Research*, ICER '18, page 41–50, New York, NY, USA, 2018. Association for Computing Machinery.

[535] James Prather, Raymond Pettit, Kayla Holcomb McMurry, Alani Peters, John Homer, Nevan Simone, and Maxine Cohen. On novices' interaction with compiler error messages: A human factors approach. In *Proceedings of the 2017 ACM Conference on International Computing Education Research*, ICER '17, page 74–82, New York, NY, USA, 2017. Association for Computing Machinery.

[536] Marc Prensky. In digital games for education, complexity matters. *Educational Technology*, 45(4):22–28, 2005.

[537] Thomas W. Price, Zhongxiu Liu, Veronica Catete, and Tiffany Barnes. Factors influencing students' help-seeking behavior while programming with human and computer tutors. *ICER 2017 - Proceedings of the 2017 ACM Conference on International Computing Education Research*, pages 127–135, 2017.

[538] Carolyn Prince, Erica Ellis, Michael T Brannick, and Eduardo Salas. Measurement of team situation awareness in low experience level aviators. *The International Journal of Aviation Psychology*, 17(1):41–57, 2007.

[539] Aare Puussaar, Adrian K Clear, and Peter Wright. Enhancing personal informatics through social sensemaking. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 6936–6942, 2017.

[540] Minna Puustinen and Lea Pulkkinen. Models of self-regulated learning: A review. *Scandinavian journal of educational research*, 45(3):269–286, 2001.

[541] Keith Quille, Natalie Culligan, and Susan Bergin. Insights on Gender Differences in CS1. In *Proceedings of the 2017 ACM Conference on Innovation and Technology in Computer Science Education - ITiCSE '17*, volume Part F1286, pages 263–268, New York, New York, USA, 2017. ACM Press.

[542] Keith Quille, Natalie Culligan, and Susan Bergin. Insights on gender differences in cs1: A multi-institutional, multi-variate study. In *Proceedings of the 2017 ACM Conference on Innovation and Technology in Computer Science Education*, ITiCSE '17, page 263–268, New York, NY, USA, 2017. ACM.

[543] M Afzalur Rahim, Clement Psenicka, Panagiotis Polychroniou, Jing-Hua Zhao, Chun-Sheng Yu, Kawai Anita Chan, Kwok Wai Yee Susana, Maria G Alves, Chang-Won Lee, Sahidur Ralunan, et al. A model of emotional intelligence and conflict management strategies: A study in seven countries. *The International journal of organizational analysis*, 2002.

[544] Daniel Railsback and Nicholas Caporusso. Investigating the human factors in esports performance. In *International Conference on Applied Human Factors and Ergonomics*, pages 325–334. Springer, 2018.

[545] Rathika Rajaravivarma. A games-based approach for teaching the introductory programming course. *ACM SIGCSE Bulletin*, 37(4):98, Dec 2005.

[546] Mladen Raković, Matthew L Bernacki, Jeffrey A Greene, Robert D Plumley, Kelly A Hogan, Kathleen M Gates, and Abigail T Panter. Examining the crit-

ical role of evaluation and adaptation in self-regulated learning. *Contemporary Educational Psychology*, 68:102027, 2022.

[547] Vennila Ramalingam and Susan Wiedenbeck. Development and validation of scores on a computer programming self-efficacy scale and group analyses of novice programmer self-efficacy. *Journal of Educational Computing Research*, 19(4):367–381, 1998.

[548] Amon Rapp. Gamification for self-tracking: From world of warcraft to the design of personal informatics systems. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–15, 2018.

[549] Meg J. Ray, Maya Israel, Chung Eun Lee, and Virginie Do. A cross-Case analysis of instructional strategies to support participation of K-8 students with disabilities in CS for all. *SIGCSE 2018 - Proceedings of the 49th ACM Technical Symposium on Computer Science Education*, 2018-Janua:900–905, 2018.

[550] Linda A Reddy, Timothy J Cleary, Alexander Alperin, and Arielle Verdesco. A critical review of self-regulated learning interventions for children with attention-deficit hyperactivity disorder. *Psychology in the Schools*, 55(6):609–628, 2018.

[551] Wolfgang Reisig. *A primer in Petri net design*. Springer Science & Business Media, 2012.

[552] Jason G Reitman, Reginald Gardner, Kathryn Campbell, Alex Cho, and Constance Steinkuehler. Academic and social-emotional learning in high school esports.

[553] Lei Ren, Jin Cui, Yi Du, and Guozhong Dai. Multilevel interaction model for hierarchical tasks in information visualization. In *Proceedings of the 6th International Symposium on Visual Information Communication and Interaction*, pages 11–16, 2013.

[554] Yanyan Ren, Shriram Krishnamurthi, and Kathi Fisler. What Help Do Students Seek in TA Office Hours? In *Proceedings of the 2019 ACM Conference on International Computing Education Research - ICER '19*, pages 41–49, New York, New York, USA, 2019. ACM Press.

[555] V.G. Renumol, Dharanipragada Janakiram, and S. Jayaprakash. Identification of cognitive processes of effective and ineffective students during computer programming. *ACM Trans. Comput. Educ.*, 10(3), August 2010.

[556] Mitchel Resnick, John Maloney, Andrés Monroy-Hernández, Natalie Rusk, Evelyn Eastmond, Karen Brennan, Amon Millner, Eric Rosenbaum, Jay Silver, Brian Silverman, et al. Scratch: programming for all. *Communications of the ACM*, 52(11):60–67, 2009.

[557] M Reyes. esports ecosystem report 2021: The key industry companies and trends growing the esports market which is on track to surpass $1.5 b by 2023. business insider. 2021.

[558] Valentin Riemer and Claudia Schrader. Playing to learn or to win? the role of students' competition preference on self-monitoring and learning outcome when learning with a serious game. *Interactive Learning Environments*, 30(10):1838–1850, 2022.

[559] Emma Riese, Olle Bälter, Björn Hedin, and Viggo Kann. Programme integrating courses fighting to get engineers to reflect on non-technical topics. *Annual Conference on Innovation and Technology in Computer Science Education, ITiCSE*, pages 133–139, 2019.

[560] Bernard Rimé. Emotion elicits the social sharing of emotion: Theory and empirical review. *Emotion review*, 1(1):60–85, 2009.

[561] Susan E Rivers, Marc A Brackett, Nicole A Katulak, and Peter Salovey. Regulat-

ing anger and sadness: An exploration of discrete emotions in emotion regulation. *Journal of Happiness Studies*, 8:393–427, 2007.

[562] Raquel Breejon Robinson, Ricardo Rheeder, Madison Klarkowski, and Regan L Mandryk. "chat has no chill": A novel physiological interaction for engaging live streaming audiences. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, New York, NY, USA, 2022. Association for Computing Machinery.

[563] Sara Rodrigues and Ana Figueiras. There and then: interacting with spatio-temporal visualization.

[564] Michael A Rosen, Wendy L Bedwell, Jessica L Wildman, Barbara A Fritzsche, Eduardo Salas, and C Shawn Burke. Managing adaptive performance in teams: Guiding principles and behavioral markers for measurement. *Human resource management review*, 21(2):107–122, 2011.

[565] Howard Rosenbaum and Pnina Shachaf. A structuration approach to online communities of practice: The case of q&a communities. *Journal of the American Society for Information Science and Technology*, 61(9):1933–1944, 2010.

[566] Mark S Rosenbaum, Mauricio Losada Otalora, and Germán Contreras Ramírez. How to create a realistic customer journey map. *Business horizons*, 60(1):143–150, 2017.

[567] Jonathan Rowe, Bradford Mott, Scott McQuiggan, Jennifer Robison, Sunyoung Lee, and James Lester. Crystal island: A narrative-centered learning environment for eighth grade microbiology. In *workshop on intelligent educational games at the 14th international conference on artificial intelligence in education, Brighton, UK*, pages 11–20, 2009.

[568] Alexander Ruf, Andreas Mühling, and Peter Hubwieser. Scratch vs. karel: Impact on learning outcomes and motivation. In *Proceedings of the 9th Workshop in*

Primary and Secondary Computing Education, WiPSCE '14, page 50–59, New York, NY, USA, 2014. Association for Computing Machinery.

[569] Marco Rüth and Kai Kaspar. Commercial video games in school teaching: two mixed methods case studies on students' reflection processes. *Frontiers in psychology*, 11:3802, 2020.

[570] Richard Ryan and James Connell. Perceived locus of causality and internalization: Examining reasons for acting in two domains. *Journal of Personality and Social Psychology*, 57(5):749–761, 1989.

[571] Jennifer L Sabourin, Lucy R Shores, Bradford W Mott, and James C Lester. Understanding and predicting student self-regulated learning strategies in game-based learning environments. *International Journal of Artificial Intelligence in Education*, 23(1):94–114, 2013.

[572] Dominik Sacha, Andreas Stoffel, Florian Stoffel, Bum Chul Kwon, Geoffrey Ellis, and Daniel A Keim. Knowledge generation model for visual analytics. *IEEE transactions on visualization and computer graphics*, 20(12):1604–1613, 2014.

[573] Jean Salac, Cathy Thomas, Chloe Butler, and Diana Franklin. Supporting diverse learners in k-8 computational thinking with tipp&see. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education*, SIGCSE '21, page 246–252, New York, NY, USA, 2021. Association for Computing Machinery.

[574] Johnny Saldaña. *The coding manual for qualitative researchers*. SAGE Publications Limited, 2021.

[575] Wojciech Samek, Thomas Wiegand, and Klaus-Robert Müller. Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models. *arXiv preprint arXiv:1708.08296*, 2017.

[576] Kari Sandouka. Interactive visualizations: A literature review. *MWAIS 2019 Proceedings*, page 8, 2019.

[577] Phil Sands and Aman Yadav. Self-regulation for high school learners in a mooc computer science course. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*, SIGCSE '20, page 845–851, New York, NY, USA, 2020. Association for Computing Machinery.

[578] Jose Luis Santos, Katrien Verbert, and Erik Duval. Empowering students to reflect on their activity with stepup!: Two case studies with engineering students. In *ARTEL@ EC-TEL*, pages 73–86. Citeseer, 2012.

[579] Jose Luis Santos, Katrien Verbert, Sten Govaerts, and Erik Duval. Addressing learner issues with stepup! an evaluation. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, pages 14–22, 2013.

[580] Stanley Schachter. The psychology of affiliation: Experimental studies of the sources of gregariousness. 1959.

[581] August-Wilhelm Scheer, Oliver Thomas, and Otmar Adam. Process modeling using event-driven process chains., 2005.

[582] Walter Schneider and Richard M Shiffrin. Controlled and automatic human information processing: I. detection, search, and attention. *Psychological review*, 84(1):1, 1977.

[583] Marlene Schommer. Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology*, 82(3):498–504, 1990.

[584] Donald A Schon. *The reflective practitioner: How professionals think in action*, volume 5126. Basic books, 1984.

[585] Gregory Schraw. Promoting general metacognitive awareness. *Instructional science*, 26(1):113–125, Mar 1998.

[586] Gregory Schraw and Rayne Sperling Dennison. Assessing metacognitive awareness. *Contemporary Educational Psychology*, 19(4):460–475, Oct 1994.

[587] Carsten. Schulte, ACM Digital Library., and ACM. Special Interest Group on Computer Science Education. *Proceedings of the 10th Koli Calling International Conference on Computing Education Research.* ACM, 2010.

[588] Dale H Schunk. Goal setting and self-efficacy during self-regulated learning. *Educational psychologist*, 25(1):71–86, 1990.

[589] Dale H Schunk. Metacognition, self-regulation, and self-regulated learning: Research recommendations. *Educational Psychology Review*, 20(4):463–467, 2008.

[590] Dale H Schunk. *Learning theories an educational perspective sixth edition.* Pearson, 2012.

[591] Dale H. Schunk, Paul R. Pintrich, and Judith L. Meece. *Motivation in Education: Theory, Research, and Applications.* Pearson Merrill Prentice Hall, Upper Saddle River, NJ, 2008.

[592] Dale H Schunk and Ellen L Usher. Social cognitive theory and motivation. *The Oxford handbook of human motivation*, 2:11–26, 2012.

[593] Sarah Schwarm and Tammy VanDeGrift. Making connections. In *Proceedings of the 8th annual conference on Innovation and technology in computer science education - ITiCSE '03*, volume 8, page 65, New York, New York, USA, 2003. ACM Press.

[594] Sarah Schwarm and Tammy VanDeGrift. Making connections: Using classroom assessment to elicit students' prior knowledge and construction of concepts. In *Proceedings of the 8th Annual Conference on Innovation and Technology in Computer Science Education*, ITiCSE '03, page 65–69, New York, NY, USA, 2003. ACM.

[595] Ralf Schwarzer and Matthias Jerusalem. Generalized self-efficacy scale. In J. Weinman, S. Wright, and M. Johnston, editors, *Measures in health psychology: A user's*

portfolio. Causal and control beliefs, pages 35–37. NFER-NELSON, Windsor, UK, 1995.

[596] Beat A Schwendimann, Maria Jesus Rodriguez-Triana, Andrii Vozniuk, Luis P Prieto, Mina Shirvani Boroujeni, Adrian Holzer, Denis Gillet, and Pierre Dillenbourg. Perceiving learning at a glance: A systematic literature review of learning dashboard research. *IEEE Transactions on Learning Technologies*, 10(1):30–41, 2016.

[597] Michael James Scott and Gheorghita Ghinea. Measuring enrichment. In *Proceedings of the tenth annual conference on International computing education research - ICER '14*, pages 123–130, New York, New York, USA, 2014. ACM Press.

[598] Kamran Sedig, Paul Parsons, and Alex Babanski. Towards a characterization of interactivity in visual analytics. *J. Multim. Process. Technol.*, 3(1):12–28, 2012.

[599] Magy Seif El-Nasr and Erica Kleinman. Data-driven game development: Ethical considerations. In *International Conference on the Foundations of Digital Games*, pages 1–10, 2020.

[600] Aleksandr Semenov, Peter Romov, Sergey Korolev, Daniil Yashkov, and Kirill Neklyudov. Performance of machine learning algorithms in predicting game outcome from drafts in dota 2. In *International Conference on Analysis of Images, Social Networks and Texts*, pages 26–37. Springer, 2016.

[601] SENPAI.gg. SENPAI.gg. https://senpai.gg/, 2021. Online; accessed April 15th 2021.

[602] Yusuf Sermet and Ibrahim Demir. Flood action vr: a virtual reality framework for disaster awareness and emergency response training. In *ACM SIGGRAPH 2019 Posters*, pages 1–2. 2019.

[603] Clifford A. Shaffer and Ayaan M. Kazerouni. The impact of programming project milestones on procrastination, project outcomes, and course outcomes: A quasi-

experimental study in a third-year data structures course. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education*, SIGCSE '21, page 907–913, New York, NY, USA, 2021. Association for Computing Machinery.

[604] David Williamson Shaffer and James Paul Gee. *How computer games help children learn.* Springer, 2006.

[605] Sadia Sharmin, Daniel Zingaro, Lisa Zhang, and Clare Brett. Impact of Open-Ended Assignments on Student Self-Efficacy in CS1. In *Proceedings of the ACM Conference on Global Computing Education - CompEd '19*, number 1, pages 215–221, New York, New York, USA, 2019. ACM Press.

[606] Judy Sheard, S. Simon, Margaret Hamilton, and Jan Lönnberg. Analysis of research into the teaching and learning of programming. In *Proceedings of the fifth international workshop on Computing education research workshop - ICER '09*, page 93, New York, New York, USA, 2009. ACM Press.

[607] Duane F. Shell, Jenefer Husman, D. M. Droesch, Indira Nath, N. Wall, and Jeannine E. Turner. Project circle: First year evaluation report (grant #r215d30195). Technical report, U.S. Department of Education: Secretary's Fund for Innovation in Education and University of Texas, College of Education, Learning Technology Center. Washington, DC, 1995.

[608] Duane F. Shell, Jenefer Husman, Jeannine E. Turner, Deborah M. Cliffel, Indira Nath, and Noelle Sweany. The impact of computer supported collaborative learning communities on high school students' knowledge building, strategic learning, and perceptions of the classroom. *Journal of Educational Computing Research*, 33(3):327–349, 2005.

[609] Duane F. Shell, Leen Kiat Soh, Abraham E. Flanigan, and Markeya S. Peteranetz. Students' initial course motivation and their achievement and retention in college cs1 courses. *SIGCSE 2016 - Proceedings of the 47th ACM Technical Symposium on Computing Science Education*, pages 639–644, 2016.

[610] Murtuza N Shergadwala and Magy Seif El-Nasr. Esports agents with a theory of mind: Towards better engagement, education, and engineering. *arXiv preprint arXiv:2103.04940*, 2021.

[611] Lei Shi and Alexandra I Cristea. Learners thrive using multifaceted open social learner modeling. *IEEE MultiMedia*, 23(1):36–47, 2015.

[612] Shneiderman. Direct manipulation: A step beyond programming languages. *Computer*, 16(8):57–69, Aug 1983.

[613] Ben Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *Proceedings 1996 IEEE symposium on visual languages*, pages 336–343. IEEE, 1996.

[614] Thomas J Shuell. Cognitive conceptions of learning. *Review of educational research*, 56(4):411–436, 1986.

[615] Shuhaida Shuhidan, Margaret Hamilton, and Daryl D'Souza. A taxonomic study of novice programming summative assessment. *Conferences in Research and Practice in Information Technology Series*, 95(Ace):147–156, 2009.

[616] Wendi Sierra. Aubrey anable, playing with feelings: Video games and affect. 2018.

[617] David Silva, Matti Tedre, and Mikko Apiola. Pedagogy of 1:1 computing in Colombia. In *Proceedings of the 13th Koli Calling International Conference on Computing Education Research - Koli Calling '13*, pages 179–187, New York, New York, USA, 2013. ACM Press.

[618] Leonardo Silva. Fostering programming students regulation of learning using a computer-based learning environment. In *2021 International Symposium on Computers in Education (SIIE)*, pages 1–5, Piscataway, NJ, 2021. IEEE.

[619] Leonardo Silva, António José Mendes, Anabela Gomes, and Gabriel Fortes Cavalcanti de Macêdo. Regulation of learning interventions in programming education:

A systematic literature review and guideline proposition. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education*, SIGCSE '21, page 647–653, New York, NY, USA, 2021. Association for Computing Machinery.

[620] Leonardo S Silva. Investigating the socially shared regulation of learning in the context of programming education. In *Proceedings of the 2020 ACM Conference on Innovation and Technology in Computer Science Education*, pages 575–576, New York, NY, 2020. Association for Computing Machinery.

[621] ME Sime, AT Arblaster, and TR Green. Structuring the programmer's task. *Journal of Occupational Psychology*, 1977.

[622] Steven Simmons, Betsy DiSalvo, and Mark Guzdial. Using game development to reveal programming competency. *Foundations of Digital Games 2012, FDG 2012 - Conference Program*, pages 89–96, 2012.

[623] Beth Simon, Sarah Esper, Leo Porter, and Quintin Cutts. Student experience in a student-centered peer instruction classroom. In *Proceedings of the ninth annual international ACM conference on International computing education research - ICER '13*, page 129, New York, New York, USA, 2013. ACM Press.

[624] Traci Sitzmann and Katherine Ely. A meta-analysis of self-regulated learning in work-related training and educational attainment: what we know and where we need to go. *Psychological bulletin*, 137(3):421, 2011.

[625] Max Sjöblom and Juho Hamari. Why do people watch others play video games? an empirical study on the motivations of twitch users. *Computers in human behavior*, 75:985–996, 2017.

[626] Aaron J. Smith, Kristy Elizabeth Boyer, Jeffrey Forbes, Sarah Heckman, and Ketan Mayer-Patel. My Digital Hand: A Tool for Scaling Up One-to-One Peer Teaching in Support of Computer Science Learning. In *Proceedings of the 2017*

ACM SIGCSE Technical Symposium on Computer Science Education - SIGCSE '17, pages 549–554, New York, New York, USA, 2017. ACM Press.

[627] Kimberly A Smith-Jentsch, Janis A Cannon-Bowers, Scott I Tannenbaum, and Eduardo Salas. Guided team self-correction: Impacts on team mental models, processes, and effectiveness. *Small Group Research*, 39(3):303–327, 2008.

[628] Alison Smith-Renner, Ron Fan, Melissa Birchfield, Tongshuang Wu, Jordan Boyd-Graber, Daniel S Weld, and Leah Findlater. No explainability without accountability: An empirical study of explanations and feedback in interactive ml. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2020.

[629] Raechel N. Soicher and Regan A. R. Gurung. Do exam wrappers increase metacognition and performance? a single course intervention. *Psychology Learning & Teaching*, 16(1):64–73, 2017.

[630] Sibel Somyürek, Peter Brusilovsky, and Julio Guerra. Supporting knowledge monitoring ability: open learner modeling vs. open social learner modeling. *Research and Practice in Technology Enhanced Learning*, 15(1):1–24, 2020.

[631] Adalbert Gerald Soosai Raj, Eda Zhang, Saswati Mukherjee, Jim Williams, Richard Halverson, and Jignesh M. Patel. Effect of Native Language on Student Learning and Classroom Interaction in an Operating Systems Course. In *Proceedings of the 2019 ACM Conference on Innovation and Technology in Computer Science Education - ITiCSE '19*, pages 499–505, New York, New York, USA, 2019. ACM Press.

[632] Elsebeth Korsgaard Sorensen and Daithí Ó Murchú. Designing online learning communities of practice: A democratic perspective. *Journal of Educational Media*, 29(3):189–200, 2004.

[633] Juha Sorva. Reflections on threshold concepts in computer programming and beyond. *Proceedings of the 10th Koli Calling International Conference on Computing Education Research, Koli Calling'10*, pages 21–30, 2010.

[634] Juha Sorva. Visual program simulation in introductory programming education; visuaalinen ohjelmasimulaatio ohjelmoinnin alkeisopetuksessa, 2012.

[635] Tom Stafford, Sam Devlin, Rafet Sifa, and Anders Drachen. Exploration and skill acquisition in a major online game. In *The 39th Annual Meeting of the Cognitive Science Society (CogSci)*. York, 2017.

[636] Piers Steel. The nature of procrastination: A meta-analytic and theoretical review of quintessential self-regulatory failure. *Psychological bulletin*, 133(1):65, 2007.

[637] Ben Stephenson, Michelle Craig, Daniel Zingaro, Diane Horton, Danny Heap, and Elaine Huynh. Exam wrappers: Not a silver bullet. In *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education*, SIGCSE '17, page 573–578, New York, NY, USA, 2017. Association for Computing Machinery.

[638] Ben Stephenson, Michelle Craig, Daniel Zingaro, Diane Horton, Danny Heap, and Elaine Huynh. Exam wrappers: Not a silver bullet. In *Proceedings of the 23rd Western Canadian Conference on Computing Education*, WCCCE '18, New York, NY, USA, 2018. ACM.

[639] Robert J Sternberg et al. *Beyond IQ: A Triarchic Theory of Human Intelligence.* CUP Archive, 1985.

[640] Martin Stigmar. Peer-to-peer teaching in higher education: A critical literature review. *Mentoring & Tutoring: Partnership in Learning*, 24(2):124–136, 2016.

[641] Jeffrey A Stone. Using reflective blogs for pedagogical feedback in CS1. In *Proceedings of the 43rd ACM technical symposium on Computer Science Education - SIGCSE '12*, page 259, New York, New York, USA, 2012. ACM Press.

[642] Jeffrey A. Stone and Elinor M. Madigan. Integrating reflective writing in cs/is. *SIGCSE Bull.*, 39(2):42–45, June 2007.

[643] Jeffrey A. Stone and Elinor M. Madigan. Integrating reflective writing in CS/IS. *ACM SIGCSE Bulletin*, 39(2):42, Jun 2007.

[644] Simone Stumpf. Horses for courses: Making the case for persuasive engagement in smart systems. In *Joint Proceedings of the ACM IUI 2019 Workshops*, volume 2327. CEUR, 2019.

[645] Leigh Ann Sudol-DeLyser. Expressions of abstraction: Self-explanation in code production. In *SIGCSE 2015 - Proceedings of the 46th ACM Technical Symposium on Computer Science Education*, pages 272–277. ACM, Inc, Feb 2015.

[646] Jarkko Suhonen, Janet Davies, Errol Thompson, and Kinshuk. Applications of variation theory in computing education. In *Proceedings of the Seventh Baltic Sea Conference on Computing Education Research - Volume 88*, Koli Calling '07, page 217–220, AUS, 2007. Australian Computer Society, Inc.

[647] Raja M Suleman, Riichiro Mizoguchi, and Mitsuru Ikeda. A new perspective of negotiation-based dialog to enhance metacognitive skills in the context of open learner models. *International Journal of Artificial Intelligence in Education*, 26(4):1069–1115, 2016.

[648] John Sweller. Cognitive load theory. In *Psychology of learning and motivation*, volume 55, pages 37–76. Elsevier, 2011.

[649] Laura Swettenham and Amy Whitehead. Working in esports: Developing team cohesion. *Case Studies in Sport and Exercise Psychology*, 6(1):36–44, 2022.

[650] Claudia Szabo and Katrina Falkner. Neo-Piagetian theory as a guide to curriculum analysis. *SIGCSE 2014 - Proceedings of the 45th ACM Technical Symposium on Computer Science Education*, pages 115–120, 2014.

[651] Claudia Szabo, Katrina Falkner, and Nickolas Falkner. Experiences in course design using neo-piagetian theory. In *Proceedings of the 14th Koli Calling International Conference on Computing Education Research - Koli Calling '14*, volume 2014-Novem, pages 81–90, New York, New York, USA, 2014. ACM Press.

[652] Claudia Szabo, Nickolas Falkner, Antti Knutas, and Mohsen Dorodchi. Understanding the effects of lecturer intervention on computer science student behaviour. *ITiCSE-WGR 2017 - Proceedings of the 2017 ITiCSE Conference on Working Group Reports*, 2018-Janua:105–124, 2018.

[653] Claudia Szabo, Nickolas Falkner, Andrew Petersen, Heather Bort, Kathryn Cunningham, Peter Donaldson, Arto Hellas, James Robinson, and Judy Sheard. Review and use of learning theories within computer science education research: Primer for researchers and practitioners. In *Proceedings of the WG Reports on Innovation and Technology in Comp Sci Education*, ITiCSE-WGR '19, page 89–109, NY, USA, 2019. ACM.

[654] Katherine A Tamminen and Peter RE Crocker. "i control my own emotions for the sake of the team": Emotional self-regulation and interpersonal emotion regulation among female high-performance curlers. *Psychology of sport and exercise*, 14(5):737–747, 2013.

[655] Katherine A Tamminen, Jeemin Kim, Chad Danyluck, Carolyn E McEwen, Christopher RD Wagstaff, and Svenja A Wolf. The effect of self-and interpersonal emotion regulation on athletes' anxiety and goal achievement in competition. *Psychology of Sport and Exercise*, 57:102034, 2021.

[656] Evelyn TS Tan, Katja Rogers, Lennart E Nacke, Anders Drachen, and Alex Wade. Communication sequences indicate team cohesion: A mixed-methods study of ad hoc league of legends teams. *Proceedings of the ACM on Human-Computer Interaction*, 6(CHI PLAY):1–27, 2022.

[657] Tang Tang, Roger Cooper, and Jake Kucek. Gendered esports: Predicting why men and women play and watch esports games. *Journal of Broadcasting & Electronic Media*, 65(3):336–356, 2021.

[658] Wanyi Tang et al. Understanding esports from the perspective of team dynamics. *The Sport Journal*, 21:1–14, 2018.

[659] Matthew E Taylor, Nicholas Carboni, Anestis Fachantidis, Ioannis Vlahavas, and Lisa Torrey. Reinforcement learning agents providing advice in complex video games. *Connection Science*, 26(1):45–63, 2014.

[660] Nicholas Taylor and Bryce Stout. Gender and the two-tiered system of collegiate esports. *Critical Studies in Media Communication*, 37(5):451–465, 2020.

[661] TL Taylor. *Play between worlds: Exploring online gaming culture.* Cambridge: The MIT Press, 2006.

[662] Matti Tedre and Peter J. Denning. The long quest for computational thinking. In *Proceedings of the 16th Koli Calling International Conference on Computing Education Research - Koli Calling '16*, pages 120–129, New York, New York, USA, 2016. ACM Press.

[663] Bruce D Temkin. Mapping the customer journey. *Forrester Research*, 3:20, 2010.

[664] Josh Tenenberg and Laurie Murphy. Knowing what i know: An investigation of undergraduate knowledge and self-knowledge of data structures. *Computer Science Education*, 15(4):297–315, 2005.

[665] Stefano Teso and Kristian Kersting. Explanatory interactive machine learning. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, pages 239–245, 2019.

[666] Renate Thies and Jan Vahrenhold. Reflections on outreach programs in CS classes: Learning objectives for "unplugged" activities. *SIGCSE'12 - Proceedings of the*

43rd ACM Technical Symposium on Computer Science Education, pages 487–492, 2012.

[667] Jakita O. Thomas. The Computational Algorithmic Thinking (CAT) Capability Flow. In *Proceedings of the 49th ACM Technical Symposium on Computer Science Education - SIGCSE '18*, volume 2018-Janua, pages 149–154, New York, New York, USA, 2018. ACM Press.

[668] Neena Thota. Repertory grid: Investigating Personal Constructs of Novice Programmers. In *Proceedings of the 11th Koli Calling International Conference on Computing Education Research - Koli Calling '11*, page 23, New York, New York, USA, 2011. ACM Press.

[669] Monique Timmers, Agneta H Fischer, and Antony SR Manstead. Gender differences in motives for regulating emotions. *Personality and social psychology bulletin*, 24(9):974–985, 1998.

[670] Ville Tirronen. Study on difficulties and misconceptions with modern type systems. In *Proceedings of the 2014 conference on Innovation & technology in computer science education - ITiCSE '14*, pages 303–308, New York, New York, USA, 2014. ACM Press.

[671] Laura Toma and Jan Vahrenhold. Self-efficacy, cognitive load, and emotional reactions in collaborative algorithms labs - a case study. In *Proceedings of the 2018 ACM Conference on International Computing Education Research*, ICER '18, page 1–10, New York, NY, USA, 2018. ACM.

[672] Laura Toma and Jan Vahrenhold. Self-Efficacy, Cognitive Load, and Emotional Reactions in Collaborative Algorithms Labs - A Case Study. In *Proceedings of the 2018 ACM Conference on International Computing Education Research - ICER '18*, pages 1–10, New York, New York, USA, Aug 2018. ACM Press.

[673] Adam J Toth, Niall Ramsbottom, Christophe Constantin, Alain Milliet, and

Mark J Campbell. The effect of expertise, training and neurostimulation on sensory-motor skill in esports. *Computers in Human Behavior*, 121:106782, 2021.

[674] Bruce W Tuckman. The development and concurrent validity of the procrastination scale. *Educational and psychological measurement*, 51(2):473–480, 1991.

[675] Selen Türkay and Sonam Adinolf. Friending to flame: How social features affect player behaviours in an online collectible card game. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–12, 2019.

[676] Selen Türkay, Jessica Formosa, Sonam Adinolf, Robert Cuthbert, and Roger Altizer. See no evil, hear no evil, speak no evil: How collegiate players define, experience and cope with toxicity. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2020.

[677] April Tyack, Peta Wyeth, and Daniel Johnson. The appeal of moba games: What makes people start, stay, and stop. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play*, pages 313–325, 2016.

[678] Bert N Uchino, John T Cacioppo, and Janice K Kiecolt-Glaser. The relationship between social support and physiological processes: a review with emphasis on underlying mechanisms and implications for health. *Psychological bulletin*, 119(3):488, 1996.

[679] Eliane RA Valiati, Marcelo S Pimenta, and Carla MDS Freitas. A taxonomy of tasks for guiding the evaluation of multidimensional visualizations. In *Proceedings of the 2006 AVI workshop on Beyond time and errors: novel evaluation methods for information visualization*, pages 1–6, 2006.

[680] Linda Valli. Listening to other voices: A description of teacher reflection in the united states. *Peabody journal of Education*, 72(1):67–88, 1997.

[681] Gerard Van den Boom, Fred Paas, and Jeroen JG Van Merrienboer. Effects of

elicited reflections combined with tutor or peer feedback on self-regulated learning and learning outcomes. *Learning and Instruction*, 17(5):532–548, 2007.

[682] Gerard Van den Boom, Fred Paas, Jeroen JG Van Merrienboer, and Tamara Van Gog. Reflection prompts and tutor feedback in a web-based learning environment: Effects on students' self-regulated learning competence. *Computers in Human Behavior*, 20(4):551–567, 2004.

[683] Wesley van den Broek, Günter Wallner, and Regina Bernhaupt. Modata–improving dota 2 experience and spectatorship through tangible gameplay visualization. In *Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts*, pages 723–730, 2019.

[684] Wil Van Der Aalst. Process mining: Overview and opportunities. *ACM Transactions on Management Information Systems (TMIS)*, 3(2):1–17, 2012.

[685] Wil Van Der Aalst. Data science in action. In *Process mining*, pages 3–23. Springer, 2016.

[686] Wil Van Der Aalst, Arya Adriansyah, Ana Karla Alves De Medeiros, Franco Arcieri, Thomas Baier, Tobias Blickle, Jagadeesh Chandra Bose, Peter Van Den Brand, Ronald Brandtjen, Joos Buijs, et al. Process mining manifesto. In *International conference on business process management*, pages 169–194. Springer, 2011.

[687] Wil Van Der Aalst, Arya Adriansyah, and Boudewijn Van Dongen. Causal nets: A modeling language tailored towards process discovery. In *International conference on concurrency theory*, pages 28–42. Springer, 2011.

[688] Gerben A van Kleef, Arik Cheshin, Lukas F Koning, and Svenja A Wolf. Emotional games: How coaches' emotional expressions shape players' emotions, inferences, and team performance. *Psychology of Sport and Exercise*, 41:1–11, 2019.

[689] Tammy VanDeGrift, Tamara Caruso, Natalie Hill, and Beth Simon. Experience report: Getting novice programmers to think about improving their software development process. In *Proceedings of the 42nd ACM Technical Symposium on Computer Science Education*, SIGCSE '11, page 493–498, New York, NY, USA, 2011. Association for Computing Machinery.

[690] Katrien Verbert, Erik Duval, Joris Klerkx, Sten Govaerts, and José Luis Santos. Learning analytics dashboard applications. *American Behavioral Scientist*, 57(10):1500–1509, 2013.

[691] Arto Vihavainen, Craig S. Miller, and Amber Settle. Benefits of self-explanation in introductory programming. In *Proceedings of the 46th ACM Technical Symposium on Computer Science Education*, SIGCSE '15, page 284–289, New York, NY, USA, 2015. Association for Computing Machinery.

[692] Arto Vihavainen, Craig S. Miller, and Amber Settle. Benefits of self-explanation in introductory programming. In *Proceedings of the 46th ACM Technical Symposium on Computer Science Education*, SIGCSE '15, page 284–289, New York, NY, USA, 2015. Association for Computing Machinery.

[693] Jennifer Villareale, Colan F. Biemer, Magy Seif El-Nasr, and Jichen Zhu. Reflection in game-based learning: A survey of programming games. In *International Conference on the Foundations of Digital Games*, pages 1–9, 2020.

[694] Jennifer Villareale, Casper Harteveld, and Jichen Zhu. " i want to see how smart this ai really is": Player mental model development of an adversarial ai player. *Proceedings of the ACM on Human-Computer Interaction*, 6(CHI PLAY):1–26, 2022.

[695] Rebecca Vivian, Katrina Falkner, Nickolas Falkner, and Hamid Tarmazdi. A method to analyze computer science students' teamwork in online collaborative learning environments. *ACM Trans. Comput. Educ.*, 16(2), February 2016.

[696] Rebecca Vivian, Katrina Falkner, and Claudia Szabo. Can everybody learn to code? Computer science community perceptions about learning the fundamentals of programming. *ACM International Conference Proceeding Series*, 2014-Novem(November):41–50, 2014.

[697] Simone Volet, Mark Summers, and Joanne Thurman. High-level co-regulation in collaborative learning: How does it emerge and how is it sustained? *Learning and Instruction*, 19(2):128–143, 2009.

[698] Henry M Walker. Collaborative learning: A case study for cs1 at grinnell college and austin. In *Proceedings of the Twenty-Eighth SIGCSE Technical Symposium on Computer Science Education*, pages 209–213, New York, NY, 1997. Association for Computing Machinery.

[699] Emily Wall, Soroush Ghorashi, and Gonzalo Ramos. Using expert patterns in assisted interactive machine learning: A study in machine teaching. In *IFIP Conference on Human-Computer Interaction*, pages 578–599. Springer, 2019.

[700] Günter Wallne. Enhancing battle maps through flow graphs. In *2019 IEEE Conference on Games (CoG)*, pages 1–4. IEEE, 2019.

[701] Guenter Wallner and Simone Kriglstein. Visualizations for retrospective analysis of battles in team-based combat games: A user study. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play*, pages 22–32, 2016.

[702] Günter Wallner. Play-graph: A methodology and visualization approach for the analysis of gameplay data. In *FDG*, pages 253–260, 2013.

[703] Günter Wallner, Nour Halabi, and Pejman Mirza-Babaei. Aggregated visualization of playtesting data. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–12, 2019.

[704] Günter Wallner and Simone Kriglstein. A spatiotemporal visualization approach

for the analysis of gameplay data. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 1115–1124, 2012.

[705] Günter Wallner and Simone Kriglstein. Visualization-based analysis of gameplay data–a review of literature. *Entertainment Computing*, 4(3):143–155, 2013.

[706] Günter Wallner and Simone Kriglstein. Multivariate visualization of game metrics: An evaluation of hexbin maps. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, pages 572–584, 2020.

[707] Günter Wallner, Simone Kriglstein, Florian Gnadlinger, Michael Heiml, and Jochen Kranzer. Game user telemetry in practice: A case study. In *Proceedings of the 11th Conference on Advances in Computer Entertainment Technology*, pages 1–4, 2014.

[708] Günter Wallner, Marnix Van Wijland, Regina Bernhaupt, and Simone Kriglstein. What players want: Information needs of players on post-game visualizations. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2021.

[709] Shu-Ling Wang and Sunny SJ Lin. The application of social cognitive theory to web-based learning through netports. *British Journal of Educational Technology*, 38(4):600–612, 2007.

[710] CM Wanlin, DW Hrycaiko, GL Martin, and M Mahon. The effects of a goal-setting package on the performance of speed skaters. *Journal of Applied Sport Psychology*, 9(2):212–228, 1997.

[711] Christopher Watson, Frederick W.B. Li, and Jamie L. Godwin. No tests required: Comparing traditional and dynamic predictors of programming success. In *Proceedings of the 45th ACM Technical Symposium on Computer Science Education*, SIGCSE '14, page 469–474, New York, NY, USA, 2014. Association for Computing Machinery.

[712] Christopher Watson, Frederick W.B. Li, and Jamie L. Godwin. No tests required: Comparing traditional and dynamic predictors of programming success. In *Proceedings of the 45th ACM Technical Symposium on Computer Science Education*, SIGCSE '14, page 469–474, New York, NY, USA, 2014. Association for Computing Machinery.

[713] Dee A. B. Weikle, Michael O. Lam, and Michael S. Kirkpatrick. Automating Systems Course Unit and Integration Testing. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education - SIGCSE '19*, pages 565–570, New York, New York, USA, Feb 2019. ACM Press.

[714] Robert Weinberg, Damon Burton, David Yukelson, and Dan Weigand. Goal setting in competitive sport: An exploratory investigation of practices of collegiate athletes. *The Sport Psychologist*, 7(3):275–289, 1993.

[715] Jerod Weinman, David Jensen, and David Lopatto. Teaching Computing as Science in a Research Experience. In *Proceedings of the 46th ACM Technical Symposium on Computer Science Education - SIGCSE '15*, number Section 4, pages 24–29, New York, New York, USA, 2015. ACM Press.

[716] Katharina Weitz, Dominik Schiller, Ruben Schlagowski, Tobias Huber, and Elisabeth André. " do you trust me?" increasing user-trust by integrating virtual agents in explainable ai interaction design. In *Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents*, pages 7–9, 2019.

[717] Etienne Wenger. *Communities of practice: Learning, meaning, and identity.* Cambridge university press, 1999.

[718] Stephen A White. Introduction to bpmn. *Ibm Cooperation*, 2(0):0, 2004.

[719] Susan Wiedenbeck. Factors affecting the success of non-majors in learning to program. *Proceedings of the 1st International Computing Education Research Workshop, ICER 2005*, pages 13–24, 2005.

[720] Laurie Williams and Richard L. Upchurch. In support of student pair-programming. *SIGCSE Bulletin (ACM, Special Interest Group on Computer Science Education)*, pages 327–331, 2001.

[721] Renee M Williams, Jean Wessel, Maryan Gemus, and Elaine Foster-Seargeant. Journal writing to promote reflection by physical therapy students during clinical placements. *Physiotherapy Theory and Practice*, 18(1):5–15, 2002.

[722] Judith D. Wilson, Nathan Hoskin, and John T. Nosek. The benefits of collaboration for student programmers. *ACM SIGCSE Bulletin*, 25(1):160–164, Mar 1993.

[723] Philip H Winne. A cognitive and metacognitive analysis of self-regulated learning: Faculty of education, simon fraser university, burnaby, canada. In *Handbook of self-regulation of learning and performance*, pages 29–46. Routledge, 2011.

[724] Philip H Winne and Allyson F Hadwin. Studying as self-regulated engagement in learning. In *Metacognition in Educational Theory and Practice*, pages 277–304. Routledge, New York, NY, 1998.

[725] Christopher A Wolters and Maryam Hussain. Investigating grit and its relations with college students' self-regulated learning and academic achievement. *Metacognition and Learning*, 10(3):293–311, 2015.

[726] John Wrenn and Shriram Krishnamurthi. Executable Examples for Programming Problem Comprehension. In *Proceedings of the 2019 ACM Conference on International Computing Education Research - ICER '19*, pages 131–139, New York, New York, USA, 2019. ACM Press.

[727] Minerva Wu, Je Seok Lee, and Constance Steinkuehler. Understanding tilt in esports: A study on young league of legends players. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–9, 2021.

[728] Benjamin Xie, Dastyni Loksa, Greg L. Nelson, Matthew J. Davidson, Dongsheng Dong, Harrison Kwik, Alex Hui Tan, Leanne Hwa, Min Li, and Amy J. Ko. A theory of instruction for introductory programming skills. *Computer Science Education*, 29(2-3):205–253, 2019.

[729] Aman Yadav, Chris Mayfield, Sukanya Kannan Moudgalya, Clif Kussmaul, and Helen H. Hu. Collaborative learning, self-efficacy, and student performance in cs1 pogil. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education*, SIGCSE '21, page 775–781, New York, NY, USA, 2021. Association for Computing Machinery.

[730] Aman Yadav, Ceren Ocak, and Amber Oliver. Computational thinking and metacognition. *TechTrends*, pages 1–7, 2022.

[731] M Adil Yalçin, Niklas Elmqvist, and Benjamin B Bederson. Cognitive stages in visual data exploration. In *Proceedings of the Sixth Workshop on Beyond Time and Errors on Novel Evaluation Methods for Visualization*, pages 86–95, 2016.

[732] Lisa Yan, Annie Hu, and Chris Piech. Pensieve: Feedback on coding process for novices. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education*, SIGCSE '19, page 253–259, New York, NY, USA, 2019. Association for Computing Machinery.

[733] Qian Yang, Aaron Steinfeld, Carolyn Rosé, and John Zimmerman. Re-examining whether, why, and how human-ai interaction is uniquely difficult to design. In *Proceedings of the 2020 chi conference on human factors in computing systems*, pages 1–13, 2020.

[734] Qian Yang, Jina Suh, Nan-Chen Chen, and Gonzalo Ramos. Grounding interactive machine learning tool design in how non-experts actually build models. In *Proceedings of the 2018 Designing Interactive Systems Conference*, pages 573–584, 2018.

[735] Iman YeckehZaare, Paul Resnick, and Barbara Ericson. A Spaced, Interleaved Retrieval Practice Tool that is Motivating and Effective. In *Proceedings of the 2019 ACM Conference on International Computing Education Research - ICER '19*, pages 71–79, New York, New York, USA, Jul 2019. ACM Press.

[736] Ji Soo Yi, Youn ah Kang, John Stasko, and Julie A Jacko. Toward a deeper understanding of the role of interaction in information visualization. *IEEE transactions on visualization and computer graphics*, 13(6):1224–1231, 2007.

[737] Ji Soo Yi, Youn-ah Kang, John T Stasko, and Julie A Jacko. Understanding and characterizing insights: how do people gain insights using information visualization? In *Proceedings of the 2008 Workshop on BEyond time and errors: novel evaLuation methods for Information Visualization*, pages 1–6, 2008.

[738] Timothy T. Yuen and Min Liu. How interactive multimedia authoring transforms object-oriented thinking. In *Proceedings of the 41st ACM technical symposium on Computer science education - SIGCSE '10*, page 426, New York, New York, USA, 2010. ACM Press.

[739] Erman Yükseltürk, Adnan Yazici, Ahmet Saçan, and Özgür Kaya. Online education experiences. In *Proceedings of the fifteenth annual conference on Innovation and technology in computer science education - ITiCSE '10*, number May 1998, page 83, New York, New York, USA, 2010. ACM Press.

[740] Jamil Zaki and W Craig Williams. Interpersonal emotion regulation. *Emotion*, 13(5):803, 2013.

[741] Veronica Zammitto. Visualization techniques in video games. *Electronic Visualisation and the Arts (EVA 2008)*, pages 267–276, 2008.

[742] Weiwei Zhang, Goran Muric, and Emilio Ferrara. Individual and collective performance deteriorate in a new team: A case study of cs: Go tournaments. *arXiv preprint arXiv:2205.09693*, 2022.

[743] Yunfeng Zhang, Q Vera Liao, and Rachel KE Bellamy. Effect of confidence and explanation on accuracy and trust calibration in ai-assisted decision making. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, pages 295–305, 2020.

[744] Lanqin Zheng, Xin Li, Xuan Zhang, and Wei Sun. The effects of group metacognitive scaffolding on group metacognitive behaviors, group performance, and cognitive load in computer-supported collaborative learning. *The Internet and Higher Education*, 42:13–24, 2019.

[745] Jichen Zhu, Katelyn Alderfer, Anushay Furqan, Jessica Nebolsky, Bruce Char, Brian Smith, Jennifer Villareale, and Santiago Ontañón. Programming in game space: how to represent parallel programming concepts in an educational game. In *Proceedings of the 14th International Conference on the Foundations of Digital Games*, pages 1–10, 2019.

[746] Jichen Zhu and Magy Seif El-Nasr. Open player modeling: Empowering players through data transparency. *arXiv preprint arXiv:2110.05810*, 2021.

[747] Caroline Ziemkiewicz, Steven Gomez, and David Laidlaw. Analysis within and between graphs: Observed user strategies in immunobiology visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1655–1658, 2012.

[748] Barry J Zimmerman. A social cognitive view of self-regulated academic learning. *Journal of Educational Psychology*, 81(3):329, 1989.

[749] Barry J. Zimmerman. Self-regulated learning and academic achievement: An overview. *Educational Psychologist*, 25(1):3–17, 1990.

[750] Barry J Zimmerman. Attaining self-regulation: A social cognitive perspective. In *Handbook of self-regulation*, pages 13–39. Elsevier, 2000.

[751] Barry J. Zimmerman. Chapter 2 - attaining self-regulation: A social cognitive perspective. In Monique Boekaerts, Paul R. Pintrich, and Moshe Zeidner, editors, *Handbook of Self-Regulation*, pages 13 – 39. Academic Press, San Diego, 2000.

[752] Barry J Zimmerman. Development and adaptation of expertise: The role of self-regulatory processes and beliefs. pages 705—-722, 2006.

[753] Barry J Zimmerman and Magda Campillo. Motivating self-regulated problem solvers. *The psychology of problem solving*, 233262, 2003.

[754] Barry J Zimmerman and Anastasia Kitsantas. Developmental phases in self-regulation: Shifting from process goals to outcome goals. *Journal of educational psychology*, 89(1):29, 1997.

[755] Barry J Zimmerman and Anastasia Kitsantas. The hidden dimension of personal competence: Self-regulated learning and practice. 2005.

[756] Barry J Zimmerman and Adam R Moylan. Self-regulation: Where metacognition and motivation intersect. 2009.

[757] Barry J Zimmerman and Manuel Martinez Pons. Development of a structured interview for assessing student use of self-regulated learning strategies. *American educational research journal*, 23(4):614–628, 1986.

[758] Barry J. Zimmerman and Manuel Martinez Pons. Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal*, 23(4):614–628, 1986.

[759] Thomas Zimmermann and Nachiappan Nagappan. Predicting defects using network analysis on dependency graphs. In *Proceedings of the 30th international conference on Software engineering*, pages 531–540, 2008.

[760] Daniel Zingaro. Peer instruction contributes to self-efficacy in cs1. In *Proceedings of the 45th ACM Technical Symposium on Computer Science Education*, SIGCSE '14, page 373–378, New York, NY, USA, 2014. ACM.

[761] Daniel Zingaro. Peer Instruction contributes to self-efficacy in CS1. *SIGCSE 2014 - Proceedings of the 45th ACM Technical Symposium on Computer Science Education*, pages 373–378, 2014.

[762] Mohammad Zohaib. Dynamic difficulty adjustment (dda) in computer games: A review. *Advances in Human-Computer Interaction*, 2018, 2018.