Metadata of the chapter that will be visualized online

Chapter Title	Learning Analytics Dashboards in Educational Games	
Copyright Year	2021	
Copyright Holder	The Author(s), under exclusive license to Springer Nature Switzerland AG	
Corresponding Author	Family Name	**Rahimi**
	Particle	
	Given Name	**Seyedahmad**
	Suffix	
	Division	
	Organization/University	University of Florida
	Address	Gainesville, FL, USA
	Email	srahimi@ufl.edu
Author	Family Name	**Shute**
	Particle	
	Given Name	**Valerie**
	Suffix	
	Division	
	Organization/University	Florida State University
	Address	Tallahassee, FL, USA
	Email	vshute@fsu.edu
Abstract	Learning analytics (LA) dashboards refer to digital tools designed to help learners keep track of their progress and goals. There is growing interest and research around the topic of LA dashboards in online learning environments, with many lessons to be learned by educational game developers and researchers. However, we need more research in this area. In this chapter we addressed these issues by reviewing the theories undergirding LA dashboards, presenting recommendations that can be used when designing LA dashboards, reviewing existing LA dashboards in educational games, and, finally, walking through an example of an LA dashboard in an educational game called *Physics Playground* (*PP*). Specifically, we illustrate how *PP* uses stealth assessment to compute students' physics understanding using gameplay data and how it presents those estimates to the students in a LA dashboard we called *My Backpack*. This process is possible through an architecture that we briefly discuss in this chapter. We conclude with our plans for expanding the LA dashboard in *PP*.	
Keywords (separated by " - ")	Dashboard - Educational games - Learning analytics - Stealth assessment - STEM education	

Chapter 24
Learning Analytics Dashboards in Educational Games

Seyedahmad Rahimi and Valerie Shute

1 Introduction

Digital games, including educational games, can be suitable vehicles for assessing and improving students' knowledge, skills, and other attributes (Clark et al., 2016; Gee, 2003; Shute & Ke, 2012). For instance, Clark et al. (2016) conducted a meta-analysis to investigate the effects of playing digital games on K-16 students' learning. Results from that meta-analysis (69 studies and collectively 6868 participants) showed that digital games significantly improved students' learning compared to nongame conditions with a moderate to strong effect size. However, despite the empirical evidence for digital games being useful for students' learning, the use of educational games in classrooms is still low (Chaudy & Connolly, 2018; Papadakis, 2018). One missing piece of the puzzle could be explicitly connecting gameplay and learning and making that visible for various stakeholders (e.g., students, teachers, parents) (Alonso-Fernández et al., 2019; Calvo-Morata et al., 2018; Chaudy & Connolly, 2018). Such visual representations of gameplay and learning are important parts of learning analytics (LA) dashboards in educational games.

According to the Society for Learning Analytics Research, the LA field is shaped around "…the measurement, collection, analysis, and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs" (Siemens & Gasevic, 2012, p. 1). LA dashboards are useful tools—for both teachers and students—as they summarize

S. Rahimi (✉)
University of Florida, Gainesville, FL, USA
e-mail: srahimi@ufl.edu

V. Shute
Florida State University, Tallahassee, FL, USA
e-mail: vshute@fsu.edu

M. Sahin, D. Ifenthaler (eds.), *Visualizations and Dashboards for Learning Analytics*, Advances in Analytics for Learning and Teaching,
https://doi.org/10.1007/978-3-030-81222-5_24

students' complex learning-related data. There is ample research done regarding LA dashboards used in online learning platforms (e.g., MOOCs, learning management systems). However, little is known about the design and effects of LA dashboards in educational games. Our chapter addresses this issue. In this chapter, we (1) define LA dashboards and discuss who can benefit from them, (2) review the relevant literature and theories about LA dashboards in general, (3) discuss recommendations about the design of LA dashboards based on the literature, (4) present examples of LA dashboards in some educational games, (5) detail the design of a particular LA dashboard in an educational game called *Physics Playground*, and (6) conclude with suggestions for future research regarding the LA dashboard in *Physics Playground*.

1.1 What Is an LA Dashboard and Who Can Benefit from It?

LA dashboards are useful tools that include visual elements (e.g., graphs, colors, and charts) generated from students' interactions in the digital environment. The data can be presented at various grain sizes and relate to different stakeholders' needs (e.g., teachers and students). According to the literature, students can benefit from LA dashboards by allowing them to set personal goals, see progress toward their goals, obtain feedback about their learning, become motivated by receiving immediate feedback, and make decisions about what to do next (Bodily et al., 2018; Jivet et al., 2017; Schumacher & Ifenthaler, 2018; Sedrakyan et al., 2020). The type of feedback that LA dashboards provide to students can be seen as formative. Decades of research on formative feedback show that it is crucial to improve students' learning (Black & Wiliam, 1998; Shute, 2008). Through formative feedback, LA dashboards can help learners make better decisions in the learning process themselves in contrast with environments where computers make the decisions for learners (e.g., via adaptive learning environments). Such environments can help learners take ownership of and consequently improve their learning via the formative feedback within LA dashboards (Charleer et al., 2016; Shute et al., 2008; Shute et al., 2020).

In some cases, dashboards permit learners to compare their progress to other students (currently in their class or historical data). Thus, LA dashboards can either show progress relative to oneself or relative to others (i.e., intrapersonal vs. interpersonal frames of reference, respectively). Choosing an appropriate frame of reference depends on a student's particular learning goal orientation. Generally, there are two goal orientations: performance orientation which refers to norm-referenced comparisons (i.e., when students compare their performance to other students) and mastery orientation which refers to criterion-referenced comparisons (i.e., when students compare their performance against a certain level of mastery) (Dweck & Leggett, 1988). Research on various LA dashboards shows that including a norm-referenced (interpersonal) frame of reference should be used cautiously. In contrast, criterion-referenced (intrapersonal) dashboards consistently show positive impacts on students' motivation and learning (e.g., Jivet et al., 2018). We discuss these frames of reference in more detail later in this chapter.

Besides students, teachers can also benefit from LA dashboards by monitoring their students' progress and evaluating the effectiveness of the learning resources they use (i.e., learning supports and materials) and their instructional methods (e.g., Alonso-Fernández et al., 2019; Calvo-Morata et al., 2018). As a result, teachers may decide to change their teaching strategy (e.g., when they see most of their students struggling on a topic) or replace some learning resources that appear ineffective. Moreover, LA dashboards can help automate various types of feedback that teachers like to provide to their students in real-time, thus saving time and effort for teachers in large classes. Importantly, dashboard data can help teachers quickly identify and help students who are struggling and intervene accordingly.

The foregoing research relates to LA dashboards that currently exist within learning environments (e.g., MOOCs or LMSs). However, our focus is on educational games as effective environments that can also benefit from rich LA dashboards. Designing such dashboards in educational games can benefit from the research done in other learning environments. To understand where these benefits are rooted, we discuss the theories behind LA dashboards next.

1.2 Theories Behind LA Dashboards

Research on LA dashboards occurs at the intersections of various disciplines, including the learning sciences, information science, learning analytics, educational data mining, psychology, and data visualization (Schwendimann et al., 2017). Therefore, LA dashboards should be designed based on the theoretical foundations from these disciplines to achieve the optimal outcomes for students and teachers (Sedrakyan et al., 2020). The following are the most important theories related to the design of LA dashboards.

According to the literature (Jivet et al., 2017, 2018; Kim et al., 2016; Sedrakyan et al., 2020), the most common learning theory underlying LA dashboards is *self-regulated learning theory* (SRL; Zimmerman, 1990). SRL refers to the metacognitive processes and strategies that a learner adopts to maximize and optimize their learning. These strategies include planning, goal setting, organizing, self-monitoring, reflecting, and adapting at various stages of learning. A self-regulated learner is self-aware, knowledgeable, and decisive in their approach to learning (Zimmerman, 1990). Learners who are self-regulated report high levels of self-efficacy and intrinsic motivation—i.e., doing something because it is internally rewarding and satisfying (e.g., Borkowski et al., 1990). One approach to help students become self-regulated learners is to teach them about strategies they can adopt (e.g., goal setting, time management, resource management). However, Zimmerman (1990) asserts that only knowing a particular strategy is not enough for a long-lasting impact of those strategies. Instead, self-regulated learning strategies should be facilitated. LA dashboards are suitable tools that can facilitate self-regulated learning (Jivet et al., 2018; Sedrakyan et al., 2020).

Similarly, LA dashboards align with the *self-determination theory* (SDT; Black & Deci, 2000). According to SDT, people feel intrinsically motivated when they gain a perception of competence, autonomy, and relatedness. Using LA dashboards, learners can monitor their progress and strategically march toward their goals, leading them to achieve high levels of competence in the targeted skills they need. Through self-awareness about their learning progress, learners can decide what to do next and gain high levels of autonomy through various choices coupled with the high level of control available in the learning environments (e.g., educational games or MOOCs). Therefore, LA dashboards can enable, rather than inhibit, student autonomy and enhance learners' intrinsic motivation.

If LA dashboards are poorly designed, learners will not (or very seldom) use them, and thus, none of the positive effects of LA dashboards will be achieved. One reason given for not using LA dashboards is the perception that they are too cluttered, confusing, and hard to understand (Jivet et al., 2017). Theories from the fields of information and communication can help make LA dashboards easier to understand. For example, sense-making theory (Dervin, 1998) indicates that "knowledge is the sense made at a particular point in time-space by someone" (p. 36). Moreover, Weick, Sutcliffe, and Obstfeld (2005) note that "sense-making involves turning circumstances into a situation that is comprehended explicitly in words and that serves as a springboard into action" (p. 409). If the information provided to the learners through LA dashboards does not make sense to them, no proper action (e.g., working more on the skill or knowledge they lack) will occur. Another example of poor design of LA dashboards is when poor computational processes lead to information that the learners disagree with. It does not make sense to them (e.g., the learners feel competent in a given skill, but the LA dashboard shows otherwise). This discrepancy can make learners lose trust in what they see on their dashboard and stop using it (Jivet et al., 2018). Therefore, it is essential to conduct various usability studies and work closely with the target audience of LA dashboards to ensure that what is presented to learners makes sense (e.g., Bodily et al., 2018; Schumacher & Ifenthaler, 2018). The application of various theories and the associated research provides the basis for various recommendations relative to designing high-quality LA dashboards.

1.3 Recommendations for LA Dashboard Design

Although these recommendations come from LA dashboard design research within online learning environments, they may be useful for LA dashboard design in educational games, as educational games can also be considered learning environments. However, students might be more intrinsically motivated to play an educational game compared to completing an online course containing the same content. Therefore, the effects of the recommendations we propose here on students' learning should be examined when used in educational games. We discuss how each recommendation relates to educational games at the end of each part and begin with choosing the appropriate frame of reference.

1.3.1 Choose the Frame of Reference Thoughtfully

Research has shown that learners need at least one of the following frames of reference to be able to interpret the LA they see in a dashboard: (1) *social*, which allows comparisons with peers (e.g., comparing one's own score with the average score of the class); (2) *achievement*, which indicates one's distance from their goals; and (3) *progress*, which allows visual self-comparison over time using their data history (Jivet et al., 2017). LA dashboards can include all three frames of reference and allow students to choose the frame in which they feel most comfortable. Providing a frame of reference by force should be done cautiously as different learners (e.g., high-achievers vs. low-achievers) may react differently to various frames of reference. Specifically, Jivet et al. (2018) reported in their literature review that low-achieving students who used a social frame of reference often became demotivated (i.e., stopped using the LA dashboard) when they saw that they were behind other students. Similarly, some high-achieving students who used the social frame of reference could become demotivated and stop working if they felt that they were better than others and did not need to do more. However, other high-achieving students found that the social frame of reference was motivating, as it provided for healthy competition. In contrast, low-achieving students who did not know how other students were doing (i.e., they were using an intrapersonal frame of reference) reported that using LA dashboards was motivating. Consequently, one recommendation from this literature is to permit learners to choose the frame of reference where they feel most comfortable and motivated. For example, students can choose to compare themselves with their classmates' average scores (i.e., social) or completely deactivate that feature and compare current performance/learning with that from an earlier stage in their learning (i.e., progress). This recommendation about using a frame of reference comes from the literature on LA dashboards in online learning environments. One could argue that using a social frame of reference could be a natural decision as games already have a competitive nature. Alternatively, since we are talking about educational games, an achievement or progress frame of reference could be helpful to students with a goal orientation, permitting them to focus on their own learning. Clearly, more research is needed to evaluate this recommendation in the context of educational games.

1.3.2 Remember that LA Is About Learning

LA dashboards use various data sources (e.g., log data from learners' interactions with a learning environment). Usually, LA dashboards visualize the data related to those interactions without emphasizing how learners are doing regarding their learning goals. These LA dashboards focus more on progress made (e.g., the number of learning modules completed in a MOOC or the number of game levels completed in an educational game) rather than learning (Gašević et al., 2015). There is a need, especially in educational games, to include psychometrically sound assessments of students' learning in educational games. The idea of LA dashboards focusing on learning relates to the open learner model (OLM) (Bull et al., 2013). By visualizing the inferences about students' learning and showing the learning analytics to the stakeholders (i.e., students and teachers), metacognitive behaviors (e.g., reflection, planning, self-awareness, self-monitoring) can be enhanced. Therefore, visualizing students' performance analytics from their interaction data is not enough—inferences about students' knowledge, skills, and other attributes are also needed. Moreover, information is needed that provides clear suggestions about how students can do better. LA dashboards in educational games, like in commercial games, tend to focus on the analytics (e.g., displaying information from log data such as minutes spent per game level). But educational game designers and researchers also need to pay close attention to the *learning* part by linking students' behavioral data to specific and pre-identified competencies (i.e., knowledge, skills, and other attributes).

1.3.3 Include "How Can I Do Better?" Functionality

Most LA dashboards focus on the "how am I performing?" question rather than "how can I do better?" (Sedrakyan et al., 2020). After a successful sense-making (or "aha!") moment when using an LA dashboard, the student will then need to take some action (Weick et al., 2005). For example, based on an analysis of a student's current understanding of Newton's first law of motion, a learning environment (e.g., an educational game) can provide behavioral instructions (e.g., "You need to solve five levels with Newton's first law as their primary concept") if the LA show that the student has not played enough Newton's first law levels. Alternatively, the game can suggest cognitive supports (e.g., "You need to watch this video explaining Newton's first law") if the player played enough targeted levels but his or her estimates are low. Providing the right formative feedback can help learners find the LA dashboard effective in which case they would use the dashboard more frequently (Kim et al., 2016).

1.3.4 Seek Feedback from Stakeholders Throughout the LA Dashboard Design Process

The main stakeholders of LA dashboards in educational games that we are focusing on in this chapter are students (or learners in general). According to the literature, conducting usability and evaluation studies when designing LA dashboards is infrequently done (Jivet et al., 2018; Sedrakyan et al., 2020). As LA dashboard designers and researchers, we need to include what learners need and expect to see in LA dashboards (Schumacher & Ifenthaler, 2018). Moreover, we need to make sure that the content in LA dashboards makes sense to the students. In this vein, some researchers have suggested including mechanisms in the learning environment to collect data about students' opinions on elements included in the LA dashboard (Jivet et al., 2018). For example, a rating system can be employed that quickly allows learners to provide feedback about various aspects of the LA dashboard in use. This recommendation can be used in educational games as well. For example, after including students throughout the design process, educational game designers and researchers can embed quick rating questions about different parts of the LA dashboard. The questions would seek input on whether students understood the information provided to them and if there were alternative formats that should be used. Next, we briefly discuss some examples of LA dashboards in educational games.

1.4 LA Dashboards in Educational Games

Most of the studies that we reviewed have focused on LA dashboards for teachers, not students (e.g., Alonso-Fernández et al., 2019; Martínez-Ortiz et al., 2019). Although some of those findings may be used to design student-focused LA dashboards in educational games, there is a gap in the literature related to studies focusing explicitly on students-aimed LA dashboards. The issue discussed earlier (i.e., collecting and reporting performance data rather than learning-related data and inferences) also exists in LA dashboards in educational games. For example, Chaudy and Connolly (2018) conducted a review on game-based learning analytics. They reported that the type of data collected in the studies they reviewed (most of them created for teachers) were time-related data, counts, game actions, scores, and player data (e.g., demographic and academic). One could argue that game performance and learning are positively related; however, we would expect to see much stronger effects on student learning if the LA in educational games were more focused on learning than performance.

We reviewed several studies that detailed the design, development, and testing of LA dashboards in educational games for students. Here we describe two of these studies. Seaton, Chang, and Graf (2019) created a game (the name of the game was not mentioned in the article) to improve students' skills (i.e., problem-solving, associative reasoning, organization and planning, and monitoring work for accuracy). This game included ten sub-games targeting the cognitive and metacognitive skills mentioned above. Each sub-game generated a score for the targeted skills in percentages based on the players' performance. There were also multiple opportunities for earning game money, badges, and points. The LA dashboard employed in this game used line graphs to visualize skill scores over time (i.e., progress), and scatterplots to visualize the relationship between performance scores and time of the day. The LA dashboard was interactive and allowed players to select a particular skill and a specific time of day or a specific sub-game to see their data visualizations. These visualizations could help players understand how their playing habits impacted their performance (e.g., using the scatterplot, the players could see how playing a sub-game at different times of the day could positively or negatively affect their performance). Also, the players could identify their strengths and weaknesses. The authors conducted a proof-of-concept evaluation using gameplay data collected over 3 months from four players. The authors claimed that the LA dashboard did provide useful information to the players. However, these results need further investigation as only four players participated in the evaluation study. Also, this evaluation study examined if what was shown to the players was meaningful and useful to them. We can argue that the LA in this study was based on performance data rather than inferences about learning. Moreover, based on what the authors provided, there were no instructions available for the students on how to interpret the line charts and scatterplots, potentially causing extraneous sense-making issues. Therefore, more rigorous studies are needed (with larger samples) to make valid conclusions about the usefulness and effectiveness of the LA dashboard in this game relative to learning.

Another example of an LA dashboard for students was developed in a game called *Selene* (Reese, 2016) about the Earth and space. In this game, players get to create their Moon by simulating an accretion process (i.e., causing collisions that can produce space debris, and then the particles would accumulate to create a massive object—a Moon). Not all types of collisions can create moons in space. Players must learn how to create collisions that include a careful balance among velocity, heat, density, and radioactivity proportions. After players learn how to create a Moon, they can then try to replicate the surface of our own Moon (created over about 4.5 billion years) by colliding meteors and flooding the Moon's surface with lava. Reese (2016) indicated that *Selene* was designed after detailed cognitive task analyses completed by subject-matter experts and then cognitive science structure mapping (Gentner, 1983). Reese claimed that "the game is the procedural analog of what is invisible inside experts' heads" (p. 236). This approach is very similar to the evidence-centered design (ECD; Mislevy et al., 2003) approach for designing an assessment. In ECD, a competency model is elaborated first (answering the question of "what is it that we want to assess?"). Then, the environment in which we can elicit evidence for the competency model is designed and developed (we will discuss ECD in more detail later in this chapter). Following this approach, students' performance data, shown on the LA dashboard, were directly linked to their mastery of the knowledge represented in the game (Reese, 2016). On *Selene's* LA dashboard, players could see their achievements (i.e., when a player completed a game level and met certain criteria), progress, and highest game score.

In both of these examples described above, players could see leader boards and compare their performance to other students (i.e., the social frame of reference), which may lead to competition rather than knowledge and skill mastery (Alonso-Fernández et al., 2018). In the next section, we discuss an example of a student-focused LA dashboard in an educational game called *Physics Playground*, which uses an achievement frame of reference and focuses on mastery, not competition.

2 Physics Playground

Physics Playground (*PP*; Shute et al., 2019a) is a 2D web-based game created to help middle- and high-school students learn Newtonian physics (e.g., Newton's laws of force and motion, energy, linear momentum, and torque). For all the game levels, the goal in this game is to direct a green ball to hit a red balloon. There are two level types: *sketching* and *manipulation* (Fig. 24.1).

To solve sketching levels, students draw simple machines (i.e., ramps, levers, pendulums, and springboards) to guide the ball to the target balloon (Fig. 24.1a). To solve manipulation levels, students interact with various sliders to change physics parameters (i.e., gravity, air resistance, mass, and bounciness of the ball) and also manipulate external forces exerted from puffers or blowers to hit the balloon—no drawing is allowed in manipulation levels (Fig. 24.1b). *PP*'s number of game levels is dynamic—we have created about 150 game levels covering nine physics competencies (Fig. 24.2). We can add game levels to the online version of *PP* at any time using the game's level editor.

2.1 Stealth Assessment

To assess students' physics understanding in real-time for each of the nine competencies, *PP* employs stealth assessment (Shute, 2011). Specifically, *PP*'s stealth assessment machinery gathers student-gameplay data in log files, automatically scores and accumulates the collected data using statistical methods (e.g., Bayesian networks), and makes real-time inferences about the current level of students' targeted competencies related to understanding Newtonian physics (see recommendation 1.3.2). Then, *PP* uses those estimates to (a) adapt game level challenges to fit a student's current competency level (for the adaptive version of the game), (b) provide appropriate learning supports to students, and (c) inform students of their progress in the game and relative to targeted physics concepts via an LA dashboard called *My Backpack* (discussed in more detail later).

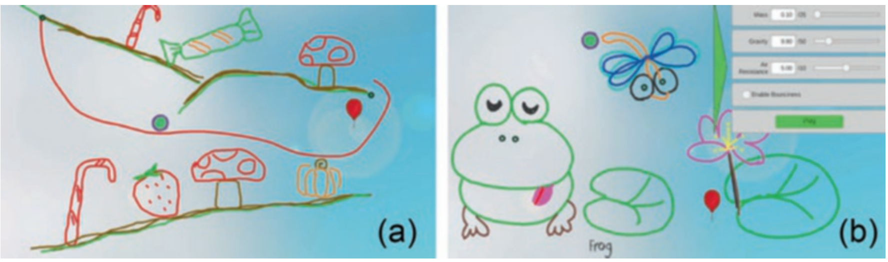

Fig. 24.1 Sketching level (**a**) and manipulation level (**b**)

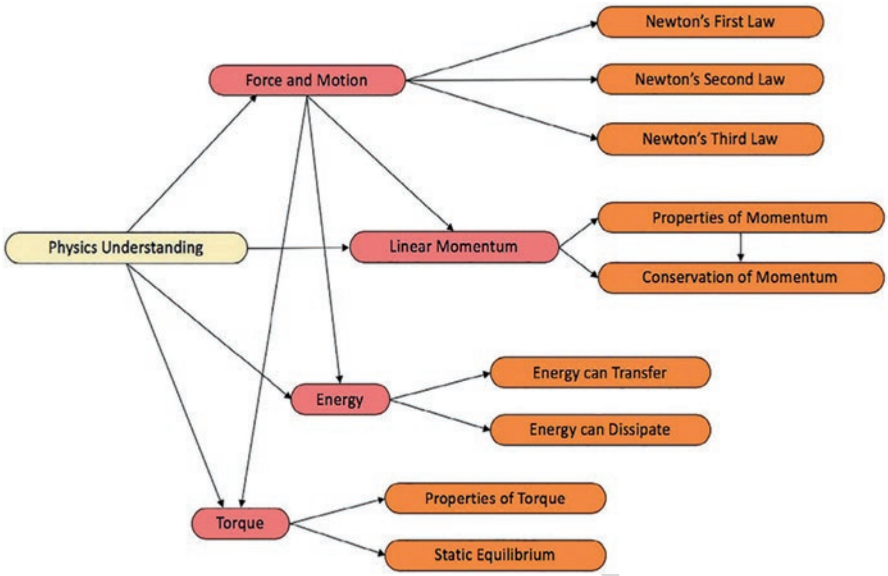

Fig. 24.2 Physics understanding competency model in *PP*

Stealth assessment is based on the evidence-centered design framework of assessment (ECD; Mislevy et al., 2003). ECD's primary purpose is to structure the collection of evidence needed to make valid claims about students' competencies (i.e., knowledge, skills, and other attributes). ECD includes a framework of conceptual and computational models that work in harmony. The three core ECD models are the following: (1) the competency model (CM), operationalizing the construct we want to assess (e.g., conceptual physics understanding) and defining the claims to be made about student competencies; (2) the evidence model (EM), automatically scoring and accumulating valid evidence (i.e., observables) of a claim about student competencies (i.e., unobservables); and (3) the task model (TM)—detailing the nature and form of the tasks (e.g., game levels) that will elicit the evidence needed for the EM.

In stealth assessment, specific gameplay behaviors are dynamically linked to the CM. As students interact with the game environment, they generate a continuous stream of data captured in the game's log files. Then, the stealth assessment tools identify and extract evidence related to the CM—in real-time—i.e., the evidence identification (EI) process. The EI's output is the input data (e.g., scores, tallies) for the evidence accumulation (EA) process, which statistically updates the claims about relevant competencies in the CM (e.g., the probability of a student being low, medium, or high on a given competency; see Almond et al., 2020 for more detail on these processes). The more evidence a student generates during gameplay, the more accurate the estimates of competency levels. As mentioned, competency-level estimates can be used for various purposes (e.g., adaptive delivery of game levels, targeted feedback, relevant learning supports, and updating the LA dashboard—*My*

Backpack). We have reported the design, development, and evaluation of various aspects of *PP* in other papers (e.g., Kuba et al., in press; Rahimi et al., 2021; Shute & Rahimi, 2020; Shute et al., 2019b, 2020). Next, we discuss the features of the LA dashboard in *PP*—*My Backpack*.

2.2 My Backpack*: PP's LA Dashboard for Students*

We designed a multipurpose dashboard in *PP* called My Backpack where students can see their progress—shown at the top part of Fig. 24.3 (i.e., the number of levels they solved, the number of gold or silver coins they collected, and the amount of money they earned). Each gold coin (given for an elegant solution for a game level) earns the student $20, and each silver coin (given for a solution that did not meet the criteria needed for a gold coin) earns $10. Students can use their game money to purchase items and customize features of the game in *PP*'s store.

In addition to showing game progress (e.g., 6 out of 22 sketching levels solved), students can monitor their level of physics understanding (Fig. 24.3) based on the current stealth assessment estimates. These estimates are for (a) each of the specific nine competencies (shown in Fig. 24.3 with the orange bar charts) and (b) their overall physics understanding (shown at the bottom of Fig. 24.3 in green). *My Backpack* also includes a store (see Fig. 24.4) where students can spend the game money they earned through gameplay to customize their game by "buying" new background music, background images, and different ball types. We designed *My Backpack* through an iterative process considering various design decisions that we mentioned in the introduction.

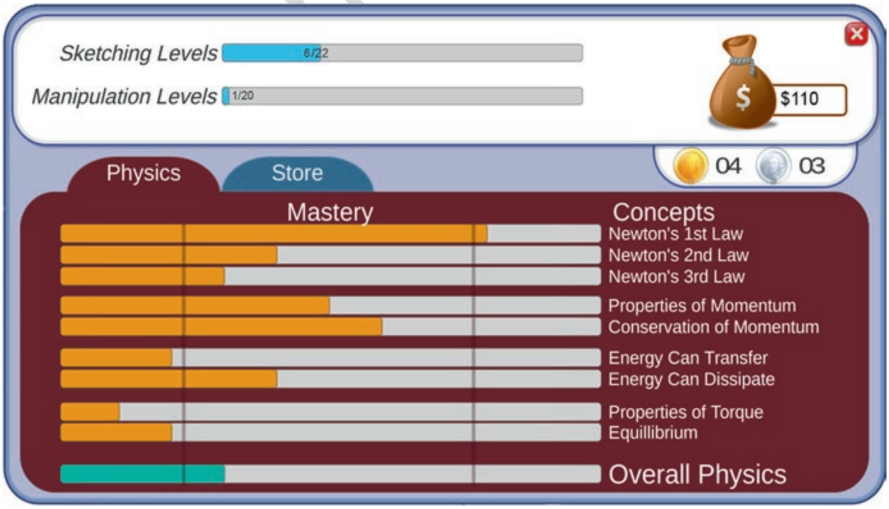

Fig. 24.3 *My Backpack*'s physics tab with indicators of student's level of competency

Fig. 24.4 Game store in *My Backpack* which includes music, background, and ball stores

2.2.1 Design Decisions Behind *My Backpack*

The first decision we needed to make was regarding the frame of reference (see recommendation 1.3.1)—social, progress, or achievement. For this version of *PP*'s LA dashboard, we decided to include an achievement (intrapersonal) frame of reference. Students can monitor their gameplay progress through the progress bars, the number of coins, and the amount of money earned (note that this is also an achievement frame of reference since students do not have access to their data history to see progress over time). Moreover, using bar charts for the nine physics competencies—the most commonly used data visualization in LA dashboards (Jivet et al., 2018; Schwendimann et al., 2017)—students can see how close they are to mastery per competency. We specifically used the word "Mastery" on top of the bar charts related to physics understanding estimates to emphasize that students should have a mastery goal (i.e., complete the bar charts) rather than a competition goal with other students. Also, because the BN estimates are dynamic (they can go up and down), students learn that if they provide negative evidence for one concept (e.g., perform poorly on a game level related to the concept that Energy Can Transfer, ECT), their level of understanding related to that particulate concept decreases. This functionality helps students build a type of mindset that they need to keep learning and doing well throughout gameplay. Consequently, they may be motivated to revisit some concepts to deepen their knowledge and achieve mastery (i.e., to complete the bar charts).

To provide various opportunities for the students to visit *My Backpack*, we made it easy to access (i.e., at the end of each game level, they would see a summary popup window indicating what money they earned in that particular level and an option to click on and visit *My Backpack*). In addition, we provided other reasons to visit *My Backpack* besides monitoring progress or achievement (i.e., we included the store that could incentivize students to use *My Backpack* more frequently). These decisions align with the principles underlying self-determination theory—i.e., providing opportunities for building competence and achieving autonomy.

We needed to translate the Bayes net estimates to a form that was understandable to students (so they can make sense of the information and then take proper actions; see recommendation 1.3.4). Consequently, we simplified the estimates. That is, instead of using three probabilities (associated with being high, medium, or low) per competency, we computed a single number (i.e., the expected a posteriori, or EAP value) ranging from −1 (low) to 1 (high) and presented that data in a bar chart (see Fig. 24.3). The EAP value for a competency is expressed as $P(\theta_{ij} = \text{High}) - P(\theta_{ij} = \text{Low})$, where θ_{ij} is the value for student i on competency j, and $[1 \times P(\text{High})] + [0 \times P(\text{Med})] + [-1 \times P(\text{Low})] = P(\text{High}) - P(\text{Low})$. Finally, to make this value even more understandable, we normalized it on a scale ranging from 0 to 1 (using this formula: $(\text{EAP} + 1) \div 2$) and showed it to the students using the orange bar charts. In our usability studies, students found *My Backpack*'s design intuitive and easy to use. Also, by providing the EAP estimates (computed via the stealth assessment machinery) to the students, we addressed the issue that LA should also be about learning—not just performance (Gašević et al., 2015). The stealth assessment process and updating of *My Backpack* is possible via *PP*'s complex architecture—discussed next.

2.3 *PP's Architecture*

A full explanation of *PP*'s architecture is outside of the scope of this chapter. Therefore, we only focus on the parts related to the stealth assessment processes and how *My Backpack* gets updated during gameplay. *PP* uses two separate servers: the *PP Server* (shown in Fig. 24.5 on the left) which hosts the game engine and the *Assessment Server* (shown in Fig. 24.5 on the right). The Assessment Server has two main components: (1) the *Dongle* component which is responsible for providing a student's prior data and their latest statistics per competency (i.e., EAPs) and (2) the assessment engine which includes two processes: evidence identification (EI) and evidence accumulation (EA).

The Dongle includes the following: (1) Proc 4 MongoDB (see Almond et al., 2020 for more details) is a filtered version of the log data, which is stored in the Learning Locker MongoDB (i.e., raw log files with much information that requires filtering; discussed below); (2) PlayerStart.php which is PHP code responsible for providing the student's previous data (i.e., levels played, coins collected, and money balance for the student) in a JSON format and interacts both with the Proc 4 MongoDB and the game engine via a POST request coming from the game engine; and (3) PlayerStats.php which is responsible for providing the student's EAPs for the nine physics competencies and overall physics understanding. These estimates are the output of the assessment engine.

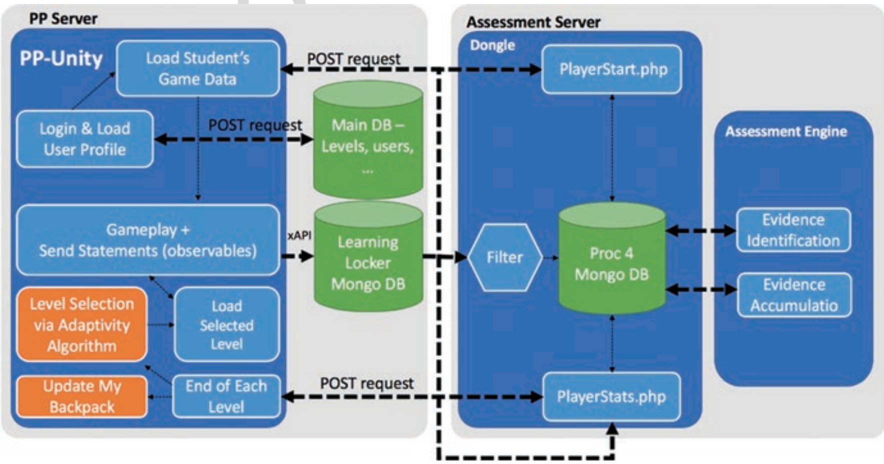

Fig. 24.5 *Physics Playground* architecture

The assessment engine has two components: (1) evidence identification (EI) whose goal is to find relevant, useful evidence in the stream of events coming from the Learning Locker and transform them into a few key observable outcomes (e.g., the coin a student received when playing a level—gold, silver, or none) and (2) evidence accumulation (EA) which is responsible for scoring the stream of observables coming from the EI process (using a Bayes net-based system) and importantly, updating the student's competency model. Using the physics understanding estimates, an adaptive algorithm in the adaptive version of *PP*—written in the game engine—selects the next level for the student (see Shute et al., 2020 for a full report about the effect of adaptivity on students' learning) and updates the student's LA output in *My Backpack*.

Learning Locker is a Learning Record Store (LRS) that stores statements generated by the xAPI-based learning activities (in this case, gaming interactions). We first specified the events or activities we needed to send to the Learning Locker. Next, we wrote various xAPI-compliant functions in the game engine when those events occurred in the game (e.g., when a level was solved and a coin was achieved). These events were sent in the form of xAPI statements to the Learning Locker. An xAPI statement consists of *actor* (i.e., user), *verb* (i.e., event), *object* (i.e., an object that the event is linked to), and *extensions* (which is a place for inserting extra data related to the event at hand—e.g., the level's name in which a particular event occurred). Learning Locker uses MongoDB, which is a document database storing data in JSON format. The Assessment Server copies and filters the raw data stored in Learning Locker—filtering out some of the xAPI metadata—for assessment purposes. Next, we discuss our plans regarding improving the LA dashboard in *PP*.

3 Future Directions for *PP*'S LA Dashboard

We envision *PP* as an engaging educational game used in classrooms (or at home) worldwide, to measure and support the learning of Newtonian physics. In one future version of *PP*, a teacher would be able to independently (without the need of possessing programming skills) create as many versions of the game with as many levels as desired for their students to play individually or collaboratively. This particular feature of *PP* (i.e., its modularity, which refers to its dynamic design capabilities) can address one of the main hurdles for using educational games in classes. That is, too often, educational games are viewed as unmodifiable black boxes that do not allow teachers to change any aspects of the game they want to use in their classes (Chaudy & Connolly, 2018). When teachers have this level of control over the game, that will instill some sense of ownership toward the game (Chaudy & Connolly, 2018), leading to more use and a higher impact on student learning.

Another logical next step with the game will entail building a dashboard for teachers to monitor their students' progress with the possibility of intervening in real time (e.g., sending feedback to students if needed). The dashboard for teachers can contain various learning analytics that can further help the teachers monitor their students' progress and learning. For example, teachers will be able to monitor progress of students individually as well as at the classroom level. Moreover, teachers could receive analytics about the effectiveness of the game resources (e.g., the efficacy of various learning supports and specific game levels). This future version of *PP* will allow teachers to dynamically add or remove any resources to and from the game based on the LA about the resources. The teacher's LA dashboard will be accessible outside of the game via an admin website to independently monitor their students' learning and progress.

To make the dashboard interpretable for teachers, we need iterative usability and experimental studies. We recommend following the suggestions from the literature about how to make LA dashboards in educational games understandable for teachers. For example, Calvo-Morata et al. (2018) suggested to (1) make LA dashboards simple rather than complex, (2) involve teachers in the dashboard design process, (3) add pop-up descriptors for complex data visualizations, and (4) add supports that can make teachers aware of undesired situations (e.g., use of alerts for statistical deviations of students from a baseline).

We also envision an advanced version of the current student dashboard in a future version of *PP*. Specifically, the student dashboard could be made to be customizable and personalized, to some extent. For example, a written interpretation/summary of the bar charts can be generated in the future to help students interpret their progress toward mastery (see recommendation 1.3.1). These features can give freedom to the students regarding their goal orientation (performance or mastery), leading to higher levels of autonomy and internal motivation (Black & Deci, 2000). To address the "how do I do better?" question (see recommendation 1.3.3), we will provide recommendations for the competencies under a certain threshold. For instance, if a student was estimated as being below some threshold relative to a concept (e.g., the EAP of ECT was less than 0.2), a pop-up menu could direct the student to either play a prescribed set of levels to enhance their knowledge about ECT or watch a targeted learning-support video about ECT before playing their next level.

Any of these future features would need to be subjected to rigorous usability and experimental testing to show relative effectiveness toward learning and performance before applied at scale. To date, testing the efficacy of the LA dashboard in *PP* has not been a primary goal. Therefore, despite following most of the recommendations about LA dashboard design, we have not collected data on the effectiveness of the LA dashboard in *PP* in terms of enhancing learning. However, we plan to conduct such studies in the future, which are intended to further help students become aware of and maximize their learning. For example, we plan to include in-game collections of usability data from students (see recommendation 1.3.4)—as suggested by Jivet et al. (2018). That is, using a simple five-star rating system, we can ask students what they think about the LA dashboard's features as they interact with each one. We will also investigate the relationship between time students spent viewing the dashboard and their motivation and learning. These investigations can shed light on how LA dashboards should be designed in educational games. In addition, in future versions of *PP*, we plan to follow the four recommendations we discussed in Sect. 1.3.

4 Conclusion

Educational games are promising tools for assessment and learning. Currently, little is known about the optimal design and effects of LA dashboards in educational games. Typically, the dashboards in educational games provide visual and textual information about learners' game performance rather than their learning. LA dashboards are tools that can help learners become aware of their learning progress and monitor their goals. There is much research around LA dashboards in online learning environments with many lessons that educational games developers and researchers can learn from. However, we need more research in this area. We addressed this issue in this chapter by reviewing theories related to LA dashboards, discussing recommendations that can be used when designing LA dashboards for educational games, reviewing LA dashboards in educational games, and finally, walking through an example of a LA dashboard in *Physics Playground*. The gap in research about LA dashboards in educational games—mainly for students—is still fairly wide. We believe that the return on investment for investigating how LA dashboards can affect students' learning in educational games will be large. Therefore, we invite our colleagues in both LA and game-based learning research areas to come together and fill this gap.

Acknowledgements This work was supported by the US National Science Foundation (award number #037988) and the Institute of Education Sciences (award number #039019). We also would like to acknowledge Russell Almond, Fengfeng Ke, Ginny Smith, Renata Kuba, Chih-Pu Dai, Curt Fulwider, Zhichun Liu, Xi Lu, and Chen Sun for helping in different phases of this project.

References

Almond, R., Shute, V., Tingir, S., & Rahimi, S. (2020). Identifying observable outcomes in game-based assessments. In *Innovative psychometric modeling and methods* (pp. 163–192). Information Age Publishing.

Alonso-Fernández, C., Cano, A. R., Calvo-Morata, A., Freire, M., Martínez-Ortiz, I., & Fernández-Manjón, B. (2019). Lessons learned applying learning analytics to assess serious games. *Computers in Human Behavior, 99*, 301–309. https://doi.org/10.1016/j.chb.2019.05.036

Alonso-Fernández, C., Pérez-Colado, I., Freire, M., Martínez-Ortiz, I., & Fernández-Manjón, B. (2018). Improving serious games analyzing learning analytics data: Lessons learned. In M. Gentile, M. Allegra, & H. Söbke (Eds.), *International conference on games and learning alliance* (pp. 287–296). Springer. https://doi.org/10.1007/978-3-030-11548-7_27

Black, A. E., & Deci, E. L. (2000). 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.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7–74. https://doi.org/10.1080/0969595980050102

Bodily, R., Kay, J., Aleven, V., Jivet, I., Davis, D., Xhakaj, F., & Verbert, K. (2018). Open learner models and learning analytics dashboards: A systematic review. In *Proceedings of the 8th* *international conference on learning analytics and knowledge* (pp. 41–50). ACM. https://doi.org/10.1145/3170358.3170409

Borkowski, J., Carr, M., Rellinger, E., & Pressley, M. (1990). Self-regulated cognition: Interdependence of metacognition, attributions, and self-esteem. In B. F. Jones & L. Idol (Eds.), *Dimensions of thinking and cognitive instruction* (pp. 53–92). Lawrence Erlbaum Associates.

Bull, S., Kickmeier-rust, M., Vatrapu, R., Johnson, M. D., Hammermueller, K., Byrne, W., … Meissl-egghart, G. (2013). Learning, learning analytics, activity visualisation and open learner model: Confusing? In *European conference on technology enhanced learning* (pp. 532–535). Springer.

Calvo-Morata, A., Alonso-Fernández, C., Freire, M., Martínez-Ortiz, I., & Fernández-Manjón, B. (2018). Making understandable game learning analytics for teachers. In G. Hancke, M. Spaniol, K. Osathanunkul, S. Unankard, & R. Klamma (Eds.), *Advances in web-based learning – ICWL 2018* (pp. 112–121). Springer. https://doi.org/10.1007/978-3-319-96565-9_11

Charleer, S., Klerkx, J., Duval, E., De Laet, T., & Verbert, K. (2016). Creating effective learning analytics dashboards: Lessons learnt. In K. Verbert, M. Sharples, & T. Klobučar (Eds.), *Adaptive and adaptable learning* (pp. 42–56). Springer. https://doi.org/10.1007/978-3-319-45153-4_4

Chaudy, Y., & Connolly, T. (2018). Specification and evaluation of an assessment engine for educational games: Empowering educators with an assessment editor and a learning analytics dashboard. *Entertainment Computing, 27*, 209–224. https://doi.org/10.1016/j.entcom.2018.07.003

Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital games, design, and learning: A systematic review and meta-analysis. *Review of Educational Research, 86*(1), 79–122. https://doi.org/10.3102/0034654315582065

Dervin, B. (1998). Sense-making theory and practice: An overview of user interests in knowledge seeking and use. *Journal of Knowledge Management, 2*(2), 36–46. https://doi.org/10.1108/13673279810249369

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*(2), 256. https://doi.org/10.1037/0033-295X.95.2.256

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

Gee, J. P. (2003). *What video games have to teach us about learning and literacy* (Vol. 1). Palgrave/Macmillan.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy*. *Cognitive Science, 7*(2), 155–170. https://doi.org/10.1207/s15516709cog0702_3

Jivet, I., Scheffel, M., Drachsler, H., & Specht, M. (2017). Awareness is not enough: Pitfalls of learning analytics dashboards in the educational practice. In É. Lavoué, H. Drachsler, K. Verbert, J. Broisin, & M. Pérez-Sanagustín (Eds.), *Data driven approaches in digital education* (pp. 82–96). Springer. https://doi.org/10.1007/978-3-319-66610-5_7

Jivet, I., Scheffel, M., Specht, M., & Drachsler, H. (2018). License to evaluate: Preparing learning analytics dashboards for educational practice. In *Proceedings of the 8th international conference on learning analytics and knowledge* (pp. 31–40). ACM. https://doi.org/10.1145/3170358.3170421

Kim, J., Jo, I.-H., & Park, Y. (2016). Effects of learning analytics dashboard: Analyzing the relations among dashboard utilization, satisfaction, and learning achievement. *Asia Pacific Education Review, 17*(1), 13–24. https://doi.org/10.1007/s12564-015-9403-8

Kuba, R., Rahimi, S., Smith, G., Shute, V., & Dai, C.-P. (in press). Using the first principles of instruction and multimedia learning principles to design and develop in-game learning support videos. *To Appear in Educational Technology Research and Development (ETRD)*.

Martínez-Ortiz, I., Freire, M., Pérez-Colado, I., Fernández-Manjón, B., & Rotaru, D. C. (2019). From heterogeneous activities to unified analytics dashboards. In *2019 IEEE global engineering education conference (EDUCON)* (pp. 1108–1113).

Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2003). Focus article: On the structure of educational assessments. *Measurement: Interdisciplinary Research & Perspective, 1*(1), 3–62. https://doi.org/10.1207/S15366359MEA0101_02

Papadakis, S. (2018). The use of computer games in classroom environment. *International Journal of Teaching and Case Studies, 9*(1), 1–25. https://doi.org/10.1504/IJTCS.2018.10011113

Rahimi, S., Shute, V., Kuba, R., Dai, C.-P., Yang, X., Smith, G., & Alonso Fernández, C. (2021). The use and effects of incentive systems on learning and performance in educational games. *Computers & Education, 165*, 104135. https://doi.org/10.1016/j.compedu.2021.104135

Reese, D. D. (2016). Dashboard effects challenge flow-learning assumption in digital instructional games. In *Emotions, technology, and digital games* (pp. 231–287). Academic. https://doi.org/10.1016/B978-0-12-801738-8.00011-7

Schumacher, C., & Ifenthaler, D. (2018). Features students really expect from learning analytics. *Computers in Human Behavior, 78*, 397–407. https://doi.org/10.1016/j.chb.2017.06.030

Schwendimann, B. A., Rodríguez-Triana, M. J., Vozniuk, A., Prieto, L. P., Boroujeni, M. S., Holzer, A., … Dillenbourg, P. (2017). Perceiving learning at a glance: A systematic literature review of learning dashboard research. *IEEE Transactions on Learning Technologies, 10*(1), 30–41. https://doi.org/10.1109/TLT.2016.2599522

Seaton, J. X., Chang, M., & Graf, S. (2019). Integrating a learning analytics dashboard in an online educational game. In A. Tlili & M. Chang (Eds.), *Data analytics approaches in educational games and gamification systems* (pp. 127–138). Springer. https://doi.org/10.1007/978-981-32-9335-9_7

Sedrakyan, G., Malmberg, J., Verbert, K., Järvelä, S., & Kirschner, P. A. (2020). Linking learning behavior analytics and learning science concepts: Designing a learning analytics dashboard for feedback to support learning regulation. *Computers in Human Behavior, 107*, 105512. https://doi.org/10.1016/j.chb.2018.05.004

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153–189. https://doi.org/10.3102/0034654307313795

Shute, V. J. (2011). Stealth assessment in computer-based games to support learning. In S. Tobias & J. D. Fletcher (Eds.), *Computer games and instruction* (pp. 503–524). Information Age Publishers.

Shute, V. J., Almond, R., & Rahimi, S. (2019a). *Physics Playground (Version 1.3)*. Tallahassee, FL. Retrieved from https://pluto.coe.fsu.edu/ppteam/pp-links/

Shute, V. J., Hansen, E. G., & Almond, R. G. (2008). You can't fatten a hog by weighing it – Or can you? Evaluating an assessment for learning system called ACED. *International Journal of Artificial Intelligence in Education, 18*(4), 289–316.

Shute, V. J., & Ke, F. (2012). Games, learning, and assessment. In D. Ifenthaler, D. Eseryel, & X. Ge (Eds.), *Assessment in game-based learning: Foundations, innovations, and perspectives* (pp. 43–58). Springer. https://doi.org/10.1007/978-1-4614-3546-4_4

Shute, V. J., & Rahimi, S. (2020). Stealth assessment of creativity in a physics video game. *Computers in Human Behavior, 116*, 106647. https://doi.org/10.1016/j.chb.2020.106647

Shute, V. J., Rahimi, S., & Lu, X. (2019b). Supporting learning in educational games: Promises and challenges. In P. Díaz, A. Ioannou, K. K. Bhagat, & J. M. Spector (Eds.), *Learning in a digital world: Perspective on interactive Technologies for Formal and Informal Education* (pp. 59–81). Springer. https://doi.org/10.1007/978-981-13-8265-9_4

Shute, V. J., Rahimi, S., Smith, G., Ke, F., Almond, R., Dai, C.-P., … Sun, C. (2020). Maximizing learning without sacrificing the fun: Stealth assessment, adaptivity and learning supports in educational games. *Journal of Computer Assisted Learning*. https://doi.org/10.1111/jcal.12473

Siemens, G., & Gasevic, D. (2012). Guest editorial-learning and knowledge analytics. *Journal of Educational Technology & Society, 15*(3), 1–2.

Weick, K. E., Sutcliffe, K. M., & Obstfeld, D. (2005). Organizing and the process of sensemaking. *Organization Science, 16*(4), 409–420. https://doi.org/10.1287/orsc.1050.0133

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist, 25*(1), 3–17. https://doi.org/10.1207/s15326985ep2501_2

Seyedahmad Rahimi is an assistant professor of educational technology in the School of Teaching and Learning at the University of Florida. Previously, he was a postdoctoral researcher at the Department of Educational Psychology and Learning Systems at Florida State University. He worked with Dr. Valerie Shute on several of her research projects. His research focuses on assessing and fostering students' 21st-century skills (focusing on creativity) and STEM-related knowledge acquisition (focusing on physics understanding) using educational games. Seyedahmad is also actively researching various aspects of educational games (e.g., game mechanics, game difficulty, cognitive and affective supports, dashboard design, and incentive systems) and how they affect students' motivation, performance, and learning. His research resulted in multiple book chapters, journal articles, and conference presentations. This chapter was written when Seyedahmad was a postdoctoral researcher at FSU.

Valerie Shute is the Mack & Effie Campbell Tyner Endowed Professor in Education in the Department of Educational Psychology and Learning Systems at Florida State University. Before coming to FSU in 2007, she was a principal research scientist at Educational Testing Service, involved with basic and applied research projects related to assessment, cognitive diagnosis, and learning from advanced instructional systems. Her general research interests hover around the design, development, and evaluation of advanced systems to support learning of knowledge, skills, and dispositions. Her research has resulted in numerous research grants, journal articles, chapters in edited books, a patent, and several recent books.

Author Query

Chapter No.: 24 0005197475

Queries	Details Required	Author's Response
AU1	Please provide better quality artwork for Figs. 24.1–24.3.	