Chapter 4

Supporting self-regulated learning with learning analytics

Jason M. Lodge¹, Ernesto Panadero², Jaclyn Broadbent³, Paula G. de Barba⁴

¹University of Queensland, Australia
²Universidad Autónoma de Madrid, Spain
³Deakin University, Australia
⁴University of Melbourne, Australia

Introduction

One of the main aims of education is ultimately to assist students to develop the capacity to learn effectively themselves. This, however, is not a simple task. Students need to develop multiple skills to be able to establish goals, monitor their progress towards those goals, correct the performance if needed, and evaluate the outcome while extracting conclusions for the next performance. They also need to become adept at finding ways to move forward when confused or otherwise reach an impasse in the learning process. Understanding how best to help students to develop these skills has been a core mission for many researchers in educational psychology and the learning sciences. These attempts have been mostly located in the self-regulated learning (SRL) field (Panadero, 2017).

For teachers, helping students to develop the skill to regulate their learning is challenging. While it is relatively straightforward to assist students with feedback about a task or process, providing feedback that will support the development of self-regulated learning is a more complex task (Hattie & Timperley, 2007). Self-regulated learning involves multiple processes including planning, monitoring, action and reflection. A significant part of this processing could be defined as forms of metacognition, alongside motivational and emotional processes, which are largely separated by some distance from obvious behavioural markers. Supporting these processes is therefore a challenging task often requiring nuanced interventions based on complex interactions between students, teachers and content (Dignath & Büttner, 2008; Panadero, 2017).

Given the complexity of the task of assisting students to develop SRL, it is perhaps unsurprising that supporting students to enhance SRL in digital learning environments is similarly complex and challenging. Despite the difficulty, with more students spending considerable amounts of time learning independently in digital environments, there is a growing need for understanding and intervening in these environments towards the development of SRL (Azevedo, Taub & Mudrick, 2017; Poitras & Lajoie, 2017). Learning analytics can be used as the basis for providing direct intervention to help students develop SRL in these environments. Learning analytics also have a role to play in assisting teachers to help students improve their capacity for SRL in these environments by providing data to teachers about how their students are progressing. In this chapter, we will discuss the elements of effective SRL and then elucidate how learning analytics may help to develop this critical set of skills in students.
Self-regulated learning has a long tradition and has become a significant topic in educational psychology over the latest two decades (Panadero, 2017). Self-regulated learning is defined as “self-generated thoughts, feelings, and actions that are planned and cyclically adapted to the attainment of personal goals” (Zimmerman, 2000 p. 14). Variables included in SRL models have systematically shown their importance and impact on students’ performance and learning (Broadbent, 2017; Broadbent & Poon, 2015; Panadero, 2017; Richardson, Abraham, & Bond, 2012). Importantly, SRL has become more prominent in recent years in computer-based learning environments. There are several key reasons for this trend which also align with broader discussions about how technology is influencing the learning process. Increasing pressure is being placed on educational institutions at all levels to deliver high-quality education to a growing number of students with greater efficiency. Preparing students for careers and lives in the complex social and economic milieu is a difficult task given the rapid pace of change in the world. Students currently completing formal education can expect to retrain and/or upskill, possibly more than once, throughout their working life. This ongoing need for education has led to the emergence of notions like 21st Century skills, which are transferrable competencies that underpin lifelong learning (Aspin, Chapman, Evans & Bagnall, 2012). We and others (e.g. de la Harpe & Radloff, 2000) suggest that SRL is critical for students to be capable of navigating this new reality. It is also apparent that SRL has a role to play when people attempt to understand complex issues. Without having adequate skill in determining how much one knows or does not know about a topic, it is easy to be misled. Some have argued that this skill is a form of critical thinking (e.g. Miller & Bartlett, 2012) or digital literacy (e.g. Eshet, 2012). Whatever the label, the reality is that students increasingly need the capability to acquire and update complex and systemic concepts and make judgements about their progress as they do so. As we have discussed, this is somewhat of a challenge given the kinds of digital environments in which students are increasingly learning.

Importantly, digital learning environments are becoming ubiquitous across all levels of education. Where once the term ‘blended learning’ was used to describe a formal educational setting in which technologies were integrated, the term could now be applied to almost all formal and informal learning. With this explosion in the use of technologies come some critical side effects that have serious implications when considering SRL. Amongst these are the increased requirement for students to be self-directed (see Hoffman & Richie, 1997). Digital learning environments are becoming increasingly adaptive and personalised, but still require students to engage in effective SRL (Greene, Moos & Azevedo, 2011) for challenges such as avoiding the temptation to multitask or access social media as they learn. Another related implication is that it is more difficult for teachers to monitor student progress. As students increasingly interact with digital environments, there is a commensurate decrease in the interaction students have with teachers. It is, therefore, more difficult for teachers working in such environments to determine when intervention is required (Arguel, Lockyer, Lipp, Lodge & Kennedy, 2017). On the other hand, some of these environments are supposed to provide means to evaluate the students and identify those who need intervention. This added complexity for teachers is compounded by continually growing class sizes. Providing personalised intervention to individual students as they need it is becoming an increasingly difficult task for teachers in the 21st Century physical and virtual classroom. A particular promise of learning analytics is the ability to help students develop their capacity in SRL to help themselves, freeing up teachers to make more targeted and nuanced interventions.
When attempting to understand what the implications of students spending time in diverse learning environments means for SRL, it is useful to consider levels of granularity in student progress. At a macro level, it is relatively easy to get a sense student achievement in their studies by looking at how they are progressing academically. In a higher education context, this represents successful subject/unit/module completion. Early applications of learning analytics, therefore, focused on this macro level of analysis (e.g. Macfadyen & Dawson, 2010). While it is relatively straightforward to see indicators that students might be struggling to self-regulate their learning at this level, it is difficult to determine why. A myriad of factors could be contributing to students not progressing as planned, which may or may not include their capacity for self-regulation in their studies.

At the opposite end of the spectrum, there has been some success in determining how students engage in self-regulated learning at the micro level in lab-based small-scale observation studies. For example, Antonietti, Colombo, and Di Nuzzo (2015) investigated self-regulation processes while students engaged in a digital learning task involving different mixes of multimedia material. The laboratory-based environment in which this study was conducted allowed for rich physiological, behavioural and self-report data to be collected (a theme we will return to shortly). In controlled environments of this kind, it is far easier to obtain multiple indicators of SRL, a luxury not afforded to digital environments “in the wild”. There is also extensive work conducted by Phil Winne with nStudy and gStudy (e.g. Perry & Winne, 2006) and by Azevedo and colleagues (e.g. Azevedo et al., 2011). The translation of the findings from laboratory environments such as these to the real world for the use of educational technologies is however complex (Lodge & Horvath, 2017). Highly controlled environments may assist in better understanding how students engage in SRL in digital and online learning, but are unlikely to provide straightforward answers to enhancing student SRL.

There is, therefore, an increase in interest and need for supporting SRL in digital environments alongside an interest in understanding what these environments mean for SRL. Helping students to navigate increasingly uncertain and contradictory channels of digital information effectively is critical for their long-term learning. It is also increasingly important to help students develop skills in cognitive tasks that are not likely to be automated in the near future. Dealing with complex, systemic knowledge is not something that machines are currently or are likely to be able to do in the very near future. However, dealing with these kinds of concepts requires an advanced capacity for SRL (Moos, 2017). Taken together, there is a distinct need for educational researchers and educators to focus on the development of SRL in digital environments, including how learning analytics may be deployed to support this imperative.

**Detecting self-regulated learning strategies**

When considering precisely what learning analytics might provide to help support the development of SRL, there are several key challenges. One of the most crucial of these challenges is the need to be able to effectively detect when critical phases of SRL are occurring, if they are developing appropriately and how to correct the student if they are not. Winne (2017) has argued that trace data is “mildly imperfect and slightly unreliable” (pg. 244) in inferring when students are engaging in SRL. However, Winne also suggests that, with care, these data can provide powerful indicators of aspects of SRL. This suggestion stands in contrast to numerous arguments that trace data has limited utility in inferring high-level cognitive processes, such as those involved in SRL (e.g. Liu, Rogers & Pardo, 2015; Lodge, Alhadad, Lewis & Gavšević, 2017; Lodge & Lewis, 2012; Rogers, 2015). So, this would suggest that an approach
built on behavioural data such as mouse clicks may need to be combined with other indicators (e.g. student performance on quizzes) and interpreted with great care to be useful in supporting SRL. In sum, there are distinct possibilities for inferring processes related to SRL from interaction behaviours as students use digital learning environments when sufficient attention is paid to the limitations of these data.

Moving away from purely behavioural data to more sophisticated approaches for detecting patterns in the learning process, there has been significant progress in several related areas in the last decade. Here we present two areas in which this progress has been most prominent. First, the use of adaptive learning environments allows for the real-time flexibility in instructional structuring and pathways through a lesson. Of particular interest for supporting SRL is that these environments incorporate aspects of student behaviour with their responses to various questions and challenges posed as they make their way through the environment. For example, platforms such as *Smart Sparrow* adapt to student responses to quiz questions as well as some aspects of their behaviour as they interact with learning tasks (see Baer, Duin, Norris & Brodnick, 2013). Where once this kind of adaptability was limited to procedural tasks such as flight simulators, it is now common for digital environments to respond and adapt to students as they learn in more complex epistemic domains (e.g. social sciences; Poitras & Lajoie, 2017). These adaptive learning environments provide examples of the kinds of data that are being used to trigger interventions to support SRL.

A second area in which significant progress has been made involve developments occurring mostly in laboratory environments using an even greater range of markers to determine how students are progressing in their learning. Azevedo, Taub and Mudrick, (2017) suggest that behavioural data can be used specifically to infer four kinds of processes: cognitive, affective, metacognitive, and motivational (CAMM). There are markers that can be detected beyond these though and potentially add to the capacity for understanding how students are progressing. These include affective markers that can determine, through facial monitoring or psychophysiological indicators, when a student is confused or frustrated, for example (Arguel et al., 2017). In a review of affect-aware technologies, Calvo and D'Mello (2010) argue that the affordances provided by these technologies will enable digital learning environments that can respond to student emotion in real time as they learn (see also D'Mello, 2017). This field is broadly referred to as affective computing (Calvo, D'Mello, Gratch, & Kappas, 2015). Along similar lines, the emerging field of multi-modal learning analytics is using similar markers to detect indicators of student progress as they learn (Azevedo et al., 2017; Ochoa, 2017). While it is unlikely that these kinds of physiological and affective instruments will see wide use in classrooms in the immediate future, the work in affective computing and multimodal learning analytics is providing a deeper understanding of the kinds of indicators that might be suggestive of higher functions such as SRL. Additionally, these type of learning environments not only measure SRL but also promote it as they include a number of instructional help elements (e.g. prompts) following earlier work (Winne’s gStudy; Hadwin, Oshige, Gress & Winne, 2010; and nStudy; Winne, Nesbit & Popowich, 2017). This combination of measurement and intervention has been analysed and discussed by Panadero, Klug & Järvelä (2016).

A crucial aspect is that these environments, and the data generated within them, should be considered in relationship to the lesson students are supposed to learn. The relationship between behavioural/trace data and the design of the lesson is an aspect that must always be taken into account when making sense of what learning analytics might mean for supporting SRL (Bakharia et al., 2016). Data cannot be considered in isolation, but must be contextualised within the instructional setting. A thorough understanding of this setting will provide deeper
insight into the markers that will help to both indicate when SRL-related processes are occurring and suggest the appropriate means of intervention. As discussed earlier in this chapter, supporting SRL in digital environments is a wicked problem. At a minimum, a systematic design approach will help to break down the elements of this issue and help to make sense of the data (Gašević, Dawson, Rogers & Gasevic, 2016). The task of determining when SRL-related processing may be going on is therefore likely to be richer when there is a clear sense of the pedagogical purpose of the lesson and all of the design elements that go into it.

There has been progress on two major fronts that may be the foundation for the effective use of learning analytics to support SRL. The first is advances in different instruments and ways of combining the data across various instruments to infer when more complex cognitive processing and affective reactions are occurring. The second is the closer integration of design and learning analytics, which allows for a more structured and contextualised means to making sense of data. Both of these lines of innovation are critical for the development of strategies and interventions that support SRL. As discussed by Panadero et al. (2016), we are now facing the third wave of SRL measurement where the evaluation and intervention co-occur, where the real potential lies in helping students develop SRL. In combination with these two trends, the systems and platforms that are becoming available for digital learning are allowing for better data collection and integration towards this goal. These trends are already evident in environments including adaptive learning environments, intelligent tutoring systems and interactive simulations.

As the mechanisms for detecting learning-related markers improve, there is also the distinct probability that learning analytics will not only help to support student SRL but also help to develop more sophisticated theoretical understandings of SRL. For example, Roll and Winne (2015) argue that the data generated through learning analytics provides insight into the stages of SRL processing. They further suggest that the kind of data generated in learning analytics research integrates well with the dominant models of SRL, particularly that of Winne and Hadwin (1998). An example of this approach is provided by Corrin, de Barba and Bakharia, (2017), who used trace data generated by students in Massive Open Online Courses (MOOCs) to explore SRL, specifically help-seeking behaviours. So, there is already significant progress in determining how and when SRL processes are occurring during student learning, but there is similarly progress in using learning analytics to contribute to a deeper theoretical understanding of SRL.

Enhancing self-regulated learning through data-driven interventions

Once SRL-related processing has been detected or predicted in a digital learning environment, there is then the question of what to do about it. There are already some promising directions for intervention approaches that help support SRL leveraging learning analytics. For example, Pardo (2017) outlines a model for data-driven feedback. A key element of this model is that the real-time feedback provided to students is specifically geared to facilitate enhanced strategies, tactics and regulation towards the pedagogical goals they are attempting to reach. Along similar lines, Timmers, Walraven, and Veldkamp (2015) drew upon behavioural trace data generated while students completed a problem-solving task in a digital environment. The data were used to provide feedback to students about their learning strategies and helped them to enhance their capacity to self-regulate their learning when later completing another problem-solving task.
What these examples of interventions have in common is that they rely on some prediction about student progress and attempt to intervene by targeting the metacognitive system. These approaches also align with the third wave of measurement of SRL described by Panadero et al. (2016) in that the measurement and intervention are deeply intertwined. The point of the intervention is to get students to stop, reflect on their progress and change their strategy in some way or another. By definition, these approaches are therefore targeting student SRL as a means of improving their development and thus their learning outcomes.

It is important to tease apart the mechanics of these intervention approaches. In the categorisation of feedback levels developed by Hattie and Timperley (2007), feedback for self-regulation is described as feedback that encourages self-monitoring, provides direction and will guide or regulate action. In this way, the intervention strategies described in this section can be seen to do exactly that. These sorts of interventions are also likely to become increasingly sophisticated as systems become more capable of predicting when best to provide feedback targeting specific aspects of monitoring, directing or acting in the learning process (Lodge, Kennedy & Hattie, 2018).

The examples we have described in this section also show how interventions based on learning analytics are moving beyond basic procedural knowledge domains to more complex domains. While it could be argued that the problem-solving process is still somewhat procedural, there is a clear trend towards environments that are increasingly able to help students come to learn more complex ideas. This trend does, however, create a significant challenge in supporting students to learn these ideas and an even greater problem when attempting to encourage them to self-regulate when learning these complex ideas. Supporting students in this way is challenging largely because students can and will adopt many different strategies to come to understand complex ideas. In research looking at how students go about completing MOOCs, for example, Milligan and Griffin (2016) found, through analysis of various forms of behavioural trace data, a wide range of strategies used by students, some effective, some less so. These findings are also evident in single sessions of instruction. When students were asked to complete a task where they created simulated stars and watched them age to learn about the effects of gravity on stellar lifecycles, their reported approaches and experiences of the module differed vastly despite the design of the lesson interface being quite structured and relatively rigid (Kennedy & Lodge, 2016). What this suggests is that any approach at supporting SRL for lessons in complex domains will either need to be robust or adaptive to this diversity or find some other such way to leverage behavioural trace data to help support SRL at critical times in the learning process.

The solution to this issue may be in the nature of the kind of cognitive processes that an intervention aimed at supporting SRL is attempting to detect and assist the student to improve. As Baker (2016) pointed out in relation to intelligent tutoring systems, the systems are actually “stupid”, and it is the humans that are smart. His argument is that adaptive intelligent tutoring systems that draw on data to facilitate learning should emphasise the enhancement of the already sophisticated processing capacity of the mind over attempting to replace it. As has undoubtedly become evident throughout this chapter, attempting to detect the kinds of high-level cognitive processes involved in SRL is complex and difficult, and still highly dependable on the researcher’s interpretations. Machines are not capable of this kind of processing. The answer to this could be to allow those processes to do what they are good at doing. The real advantages of these systems therefore comes through leveraging and enhancing human capacity rather than doing all the work for the students. The point here is that the interventions that can support SRL in digital environments can be aimed more at broadly engaging
monitoring, directing or action, nudging students towards adopting different strategies rather than attempting to provide specific feedback or advice. These kinds of interventions would, therefore, look more like personalised suggestions, hints or prompts than extensive feedback or precise guidance. In their review of feedback processes, Hattie and Timperley (2007) found that this cueing is amongst the most powerful forms of feedback, suggesting that SRL interventions could be more effective as nudges rather than shoves.

One prominent example of these kinds of approaches is provided by work on metacognitive prompts (Schworn & Renkl, 2007). Metacognitive prompts are questions or explanations that prompt the student to reflect on the success of their learning processes, their progress, and learning outcomes. For example, MetaTutor (Azevedo, Johnson, Chauncey & Burkett, 2010) uses metacognitive prompts during learning by providing feedback to enhance academic achievement. Importantly, this feedback encourages metacognitive strategies by prompting the student to correct ineffective learning strategies and replace them with new, more effective ones. The apparent impact of metacognitive prompts suggests that the intervention strategies that target SRL-related processing do not need to be either comprehensive or precise. Simply prompting students to stop and consider whether the strategies they are using in their studies are effective, may be enough to get them to monitor their progress more closely and redirect their activity or employ different strategies. Much of the work in this instance can, therefore, be left to the intelligent student and not the ‘stupid’ machine.

When considering the progress in both the means by which SRL-related processes can be detected and predicted and how it might be possible to provide interventions to advance SRL, the complexity of the task may not be a showstopper. Given that it is relatively easy to prompt or cue students into thinking more carefully about their progress, it may be more straightforward to get students to engage in SRL-related behaviours than it may seem on the surface. That being the case, the challenge for learning analytics researchers will then be to move beyond just prompting SRL, but also to help make sure that students develop more accuracy in the judgements they make about their work and their progress and to make better decisions about how best to progress (Lodge, Kennedy & Hattie, 2018). Combining this with more sophisticated and user-friendly platforms for delivering digital learning, and increasingly about complex ideas, learning analytics is already impacting on the affordances these environments provide for enhancing SRL and will continue to become better at doing so in the future (see also Lodge, 2018).

**In the classroom**

Much of the focus in this chapter has been on learning in digital environments as it is in these environments that the majority of progress in the use of learning analytics to support SRL has occurred. This may give the impression that teachers and educational researchers do not have a stake in these innovations. That, however, is far from the truth. There are short-term and long-term implications for educational researchers and teachers. In the short term, teachers will still play a critical role in directing the pedagogy that will underpin the kinds of digital environments that are being developed. While these may increasingly become able to support students to self-regulate, that is not to say that the design of the task will help students make complete sense of what they are learning. Knowledge needs to be contextualised and embodied in ways that machines, at least as they are currently conceptualised, cannot do. So, teachers will play a critical role in helping students to develop SRL no matter the sophistication of the tools. Teachers and educational researchers also need to be intimately involved in the development
of the technologies as they evolve. Teachers provide the most sophisticated approach to enhancing student SRL still. Teachers need to interpret how best to use data through the lens of learning design (Alhadad, Thompson, Knight, Lewis & Lodge, 2018). In this fashion, Bakharia et al. (2016) presented a framework on how teachers can deal with different types of data. The critical component is that the teacher is firmly at the centre.

In a similar vein, given the complexity of SRL, including how to detect related processing and intervene, there is real power in the innovations described in this chapter to support teachers. The most obvious example of this is through the use of dashboards or other mechanisms for providing teachers with information about how their students are travelling. This information also helps alleviate the “stupid machine-smart human” issue by putting the meaning-making back in the hands of the teacher. For example, students may be working through a module on a complex conceptual notion. While they do so, information is sent in real time to the teacher at the front of the class. This information could include how long students are spending on different screens, what they are clicking on or how they are responding to questions or prompts from the system. If a student gets stuck, for example, it might be flagged because they are spending more time on a single screen than expected. With this information then sent to the teacher, they are able to work with that individual student to help resolve the impasse in a far more sophisticated way than the system is likely to. Corrin and colleagues (2015) describe a tool called ‘Loop’, which provides sophisticated visualisations to teachers to allow exactly this kind of intervention. The system and the teacher working together, therefore, provide a means of targeting interventions to the students that need help.

Over the longer term, teachers will continue to play a critical role in providing all the social and embodied aspects of learning, including SRL. Technologies will become more sophisticated and able to provide personalised, adapted experiences to students in real time. This will include the ability to help support their SRL both indirectly (i.e. nudges, prompts) and directly (i.e. comprehensive, targeted suggestions specific to the student and the lesson). Despite this, students will need to operate in a social world, and SRL does not only occur in the vacuum of a digital learning environment. Teachers will continue to provide the nuanced and critical capabilities students need to enact their capacity for SRL in that social world. With less of the job of teachers to help students to acquire declarative knowledge, the important role of teachers to design digital environments and work with the data students produce while making meaning of the knowledge will become even more vital.

References


