Chapter 21: Learning Analytics for Self-Regulated Learning

Philip H. Winne

Faculty of Education, Simon Fraser University, Canada
DOI: 10.18608/hla17.021

ABSTRACT

The Winne-Hadwin (1998) model of self-regulated learning (SRL), elaborated by Winne's (2011, in press) model of cognitive operations, provides a framework for conceptualizing key issues concerning kinds of data and analyses of data for generating learning analytics about SRL. Trace data are recommended as observable indicators that support valid inferences about a learner's metacognitive monitoring and metacognitive control that constitute SRL. Characteristics of instrumentation for gathering ambient trace data via software learners can use to carry out everyday studying are described. Critical issues are discussed regarding what to trace about SRL, attributes of instrumentation for gathering ambient trace data, computational issues arising when analyzing trace and complementary data, the scheduling and delivery of learning analytics, and kinds of information to convey in learning analytics that support productive SRL.

Keywords: Grain size, metacognition, self-regulated learning (SRL), traces

Four descriptions of learning analytics are widely cited. Siemens (2010) described learning analytics as "the use of intelligent data, learner-produced data, and analysis models to discover information and social connections, and to predict and advise on learning." The website for the 1st International Conference on Learning Analytics and Knowledge posted this description: "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."[1] Educause (n.d.) defined learning analytics as "the use of data and models to predict student progress and performance, and the ability to act on that information." Building on Eckerson's (2006) framework, Elias (2011) notes "learning analytics seeks [sic] to capitalize on the modelling capacity of analytics: to predict behaviour, act on predictions, and then feed those results back into the process in order to improve the predictions over time" (p. 5).

These descriptions beg fundamental questions. What data should be gathered for input to methods that generate learning analytics? Answering this question sets boundaries on and shapes, first, approaches to computations that underlie analytics and, second, what analytics can say about phenomena. For instance, ordinal (rank) data preclude using arithmetic operations on data, such as addition or division. If data are not ordinal, A cannot be described as greater than B, nor are transitive statements valid: if A > B and B > C, then A > C.

What properties of data bear on the validity of interventions based on learning analytics developed from the data? For example, determining that a learner's age, sex, or lab group predicts outcomes offers weak grounds for intervening without other data. None of these data classes are legitimately considered a direct, proximal (i.e., sufficient) cause of outcomes. Age and sex can't be manipulated; changing lab group may be impractical (e.g., due to scheduling conflicts with other courses or a job). And, notably, prediction does not supply valid grounds for inferring causality.

Who generates data? Who receives learning analytics? Learning ecologies involve multiple actors. Authors of texts and web pages vary cues they intend to guide learners about how to study; font styles and formats (bullet lists, sidebars that translate text to graphics)

1 https://tekri.athabascau.ca/analytics/

are examples. Instructional designers and instructors augment authors' works, for example, by setting goals for learning and elaborating content. They create and recommend schedules for learning; they control most opportunities for feedback to learners. Learners study solo and often form online cliques or study groups in which they exchange views about topics, share products of learning activities (e.g., questions, notes), and form and disengage from social units. The college or university strives to improve material and cyber infrastructure wherein other actors' work unfolds. Each category of actors generates data and is a legitimate candidate to receipt of learning analytics.

What are the temporal qualities – onset, duration, and offset – of collecting data, processing it, and delivering learning analytics? Will learners receive learning analytics as they work or will they need to be reminded of context when learning analytics are delayed? Are temporal delimiters positioned elastically or rigidly across a timeline of learning? Whose model of a learning episode – the analyst's or the learner's – matters?

Finally, what are learning analytics supposed to help improve? And, what standards should be used to gauge improvement? For example, if after receiving learning analytics a learner becomes more efficient in studying but achievement does not improve, is this a benefit? Is there value in freeing time for learners to engage in activities beyond academic assignments?

In this chapter, in keeping with a focus on *self*-regulated learning, the learner is positioned as the prime actor. Other actors' activities play roles as external conditions that may vary and, perhaps, be influenced by a learner's behaviour.

Self-Regulated Learning

A framework is useful to conceptualize learning analytics for self-regulated learning (SRL). When learners self-regulate their learning, they "actively research what they do to learn and how well their goals are achieved by variations in their approaches to learning" (Winne, 2010a, p. 472). One widely cited model elaborates features of SRL as four loosely sequenced recursive phases that unfold over the timeline of a task (Winne, 2011; Winne & Hadwin, 1998).

In phase 1, a learner surveys resources and constraints the learner predicts may affect how work on a learning task proceeds, the probability that specific actions bring about particular results, and the consequences of those activities. These factors can be located externally, in the learning environment or internal to the learner. Examples of external factors include access to information available from peers or in the Internet, software tools with functions designed to support learning in various ways, and time allowed for work on a task. Examples of internal factors include knowledge and misconceptions, interest in the task or topic, or a motivational disposition to interpret slow progress as a signal of low ability or of need to apply more effort (see Winne, 1995).

Having identified resources and constraints, in phase 2 a learner sets goals and plans how to approach them. Goals are standards a product should meet. Ipsative goals compare a learner's current results to earlier results; they measure personal growth (or decline). Criterion-referenced goals measure a product in relation to a fixed profile of task features or achievements in a domain. Norm-referenced goals position a learner's product relative to a peer's or a group's. Comparisons may be framed by a learner, an instructor, or other person. It is important to note that goals can target attributes of learning processes: which process is used, effort dedicated to carrying out a process, efficiency of a process, or increasing the probability a process yields a particular product. Goals also can be set in terms of products per se and their attributes; for example, number of pages written for an essay, anxiety reduced, or thoroughness of exposition. Plans describe actions a learner intends to carry out to approach goals. Every action potentially generates multiple products. Key products include information added to knowledge, errors corrected, gaps filled or misconceptions replaced. Products can also include the learner's perceptions about rate of progress, effort spent, opportunity to explore, or prospects to impress others.

In phase 3, the learner engages with the task by enacting planned operations. Working on a task inherently generates feedback that updates the task's conditions. Feedback may originate in the learner's external environment, such as when software beeps or a peer comments on a contribution to an online discussion. Or, feedback may arise internally as a result of the learner's monitoring work flow, such as when a search query is deemed unproductive because results were not what was expected or don't satisfy the need for particular information. Modest "course corrections" may result as the learner tracks updates to conditions across the timeline of a task. It is worth explicitly noting that goals can be updated.

Phase 4 is when the learner disengages from the task as such, monitors results in one or several of phases 1 to 3, and elects to make a large-scale change. Examples might be when a learner suspends work on solving a problem and returns to studying assigned readings with a goal to repair major gaps in knowledge; or, if re-studying is not predicted to be successful, the learner asks for help from the instructor. Changes may be immediately applied to the task, reshaping

Table 21.1. SMART Cognitive Operations

Operation	Description	Sample Traces
Search	Directing attention to particular information	Opening successive bookmarks. Using a search tool.
Monitor	Comparing information presentations in terms of standards	Highlighting text (the information highlighted meets a standard, e.g., important). Selecting a previously made note for review (e.g., judgment of learning).
Assemble	Relating items of information	Tagging. Assigning two bookmarks to a titled folder.
Rehearse	Maintaining or re-instating information in working memory	Reviewing a note. Copying, then pasting.
Translate	Transforming the representation of information	Paraphrasing. Describing a graph, equation, or diagram in words.

its multivariate profile in a major way. Or, plans for change may be filed for later use, what is called "forward reaching transfer."

A 5-slot schema describes elements within each phase of SRL. A first-letter acronym, COPES (Winne, 1997, summarizes the five elements in this schema. C refers to conditions. These are features the learner perceives influence work throughout phases of the task. For example, if there are no obvious standards for monitoring a product generated in phase 3, the learner may elect to search for standards or may abandon the task as too risky. Conditions fall into two main classes, as noted earlier. Internal conditions are the learner's store of knowledge about the topic being studied and about methods for learning, plus the learner's motivational and affective views about self, the topic, and effort in this context. External conditions are factors in the surrounding environment perceived to potentially influence internal conditions or two of the other facets of COPES, operations and standards.

O in the COPES schema represents operations. First-order or primitive cognitive operations transform information in ways that cannot be further decomposed. I proposed five such operations: searching, monitoring, assembling, rehearsing, and translating; the SMART operations (Winne, 2011). Table 21.1 describes each along with examples of traces – observable behaviour – that indicate an occurrence of the operation. Second- and higher-order descriptions of cognition, such as study tactics and learning strategies, are modelled as a pattern of SMART operations (see Winne, 2010a). An example study tactic is "Highlight every sentence containing a definition." An example learning strategy is "Survey headings in an assigned reading, pose a key question about each, then, after completing the entire reading assignment, go back to answer each question as a way to test understanding."

P is the slot in the COPES schema that represents products. Operations inevitably create products, though not always intended ones. A product can be uncomplicated, such as an ordered list of British monarchs, or complex, for example, an argument about privacy risks in social media or an explanation of catalysis. E represents evaluations of a product relative to standards, S, for products. Standards for a product constitute a goal.

Two further and significant characteristics of SRL are keys to considering how learning analytics can inform and benefit learners. First, SRL is an adjustment to conditions, operations, or standards. Thus, SRL can be observed only if data are available across time. Second, learners are agents. They regulate their learning within some inflexible and some malleable constraints, the conditions under which they work. As agents, however, learners always and intrinsically have choice as they learn. A learner may think, "I did it because I had to." A valid interpretation is that the learner elected to do it because the consequences forecast for not doing it were sufficiently unappealing as to outweigh whatever cost was levied by doing it.

The COPES model identifies classes of data with which learning analytics about SRL can be developed. In the next major section, I describe four main classes of data distinguished by their origin: traces, learner history, reports, and materials studied. In the following major section, I examine computations and reporting formats for learning analytics in relation to SRL. Together, these sections describe an architecture for learning analytics designed to support learners' SRL. In a final section, I raise several challenges to designing learning analytics that support SRL.

DATA FOR LEARNING ANALYTICS ABOUT LEARNING AND SRL

As learners work, they naturally generate ambient data (sometimes called accretion data; Webb, Campbell, Schwartz, & Sechrest, 1966). Ambient data arise in the natural course of activity. For example, clicking a hyperlink to open a web resource is data about a learner's cognition and motivation – based on whatever is the present context (perhaps the title of the resource), the learner forecast information in it has sufficient value to motivate viewing it. This click is a trace, a bit of ambient data that affords relatively strong inferences about one or more cognitive, affective, metacognitive, and motivational states and processes (CAMM processes; Azevedo, Moos, Johnson, & Chauncey, 2010). I offer two further examples of traces and inferences they afford. An explicit caution is the validity of inferences grounded in trace data should always be qualified by a probability <1.00 (certainty).

Highlighting a Sentence Fragment. To select particular text for highlighting among hundreds of sentences read in a typical study session, the learner metacognitively monitors attributes of information in the text relative to standards. Standards discriminate text to be highlighted from text that should not be highlighted. The learner might monitor information for "structural" features, such as definitions or principles; or for motivational/affective features, such as interestingness or novelty. Authors often attempt to signal information that should be highlighted using font styles (e.g., italics) or phrasing: "It is interesting that…" A highlight also traces that the learner plans to review highlighted text. Why else would the learner permanently mark selected text?

Reviewing a Note. Before reviewing a particular note, the learner engages in metacognitively monitoring what can be recalled about or what is understood about particular information. The learner chooses to review when recall is judged sufficiently inadequate perhaps because it is inaccurate, incomplete, or unclear. Searching for and re-viewing a particular note traces motivation to repair whatever problem the learner perceived. If the learner highlights information in the reviewed note, that identifies which particular information the learner had monitored and judged inadequate.

Four features describe ideal trace data gathered for generating learning analytics to support SRL. First, the sampling proportion of operations the learner performs while learning is large. Ideally, but not realistically, every operation throughout a learning episode is traced. Second, information operated on is identifiable. Third, traces are time stamped. Fourth, the product(s) of operations is (are) recorded. Data having this 4-tuple structure would permit an ideal playback machine to read data about a learning episode and output a perfectly mirrored rendition of every learning event and its result(s). With 4-tuple trace data, raw material is available to generate rich learning analytics.

In reality, every trace datum is at least mildly imperfect and slightly unreliable. For example, a highlight traces a monitoring operation and generates a product – the mark plus the content marked. At a future time, this marked content is a condition that facilitates focused review. What may not be clearly revealed by a highlight is the standard(s) the learner used to select that content. Better designed traces can reduce this gap. If learners are invited to tag content they highlight – interesting, important, unclear, project1, tellMike, – each tag the learner applies exposes a standard used to metacognitively monitor the information highlighted. In some cases, a tag reveals a stronger signal about a plan – e.g., use this content in project1, in the next chat tell Mike about this content.

Learner History

Instruments for recording traces to mirror the history of a learner's activities are available in at least three environments: paper systems, learning management systems, and systems that offer learners tools for studying "on the fly."

Paper Systems. In a paper-based learning environment, some examples of traces include content highlighted, notes, marginalia such as !, ?, and √ added to the whitespace of textbook pages, a pile of books or papers stacked in order of use (e.g., the topmost was most recently used), and post-it tabs of various colours attached to pages in a notebook.

Consider the ? symbol a learner may write in the margin of a textbook page. This symbol traces that the learner metacognitively monitored the meaning of content and judged it confusing or lacking information needed to fully grasp it. A further inference is available. Why would the learner spend effort to write the ? symbol in the margin? Content could be judged confusing or incomplete without recording a symbol. Odds are the learner is motivated to and plans to repair this gap, and return to context surrounding the text to improve understanding.

While tracing in a paper-based environment is easy for learners to do, it is hugely labour intensive to gather and prepare paper-bound trace data for input to methods that compute learning analytics. In software-supported environments, this burden is greatly eased.

Learning Management Systems. Today's learning management systems seamlessly record several time-stamped traces of learners' work, such as logging in and out, resources viewed or downloaded, assignments uploaded, quiz items attempted, and forum posts to anyone or to particular peers. By adding some simple interface features, goals can be inferred. For example, clicking a button labelled "practice test" traces a learner's judgment that knowledge is lacking or certitude is below a threshold of confidence. Aggregate trace data can support inferences about 1) learners' preferred work schedules that mildly support inferences about procrastination, 2) which resources are judged more relevant or appealing, 3) motivation to calibrate judgments of learning and efficacy, and 4) value attributed to contributing, acquiring, or clarifying by exchanging information with peers.

Traces gathered across the time stream can mark when learners first study a resource, if and when they review a resource, if and when they choose to self test, and when they take a test for marks. Coupled with other data about factors such as credit hours completed or the characteristics of peers with whom information is exchanged, traces like these provide raw material for building models about how learners self-regulate the study-review-practice-test cycle (Arnold & Pistilli, 2012; Delaney, Verkoeijen, & Spirgel, 2010; Dunlosky & Rawson, 2015).

When instructors or institutions require students to use a learning management system, ambient data are generated in the course of everyday use of the system. Costs incurred to collect trace data and prepare them for input to computations that generate learning analytics are slight.

Most learning management systems lack precision in traces with respect to tracking operations learners carry out as they study or review, and which particular information they study and review. A time-stamped trace that a resource was downloaded provides no information about whether the learner studied that content, not to mention how the learner studied it.

Software Tools for Studying. Winne and Baker (2013) nominated a triumvirate of motivation, metacognition and SRL as "raw material for engineering the bulk of an account about *why* and *how* learners develop knowledge, beliefs, attitudes and interests" (p. 1). They noted three challenges to research on improving learning outcomes by mining trace data about these factors: operationalizing indicators, gathering data that trace these constructs and filtering noise that obscures signals about the constructs (see also Roll & Winne, 2015a).

Operationalizing indicators – traces of COPES – calls for software developers to exercise imagination in designing interfaces that optimize opportunities for gathering trace data while supporting experimentation about learning and without enforcing new or perturbing a learner's preferred work habits. Table 21.2 presents illustrations of opportunities to gather trace data in a context where the learner uses software tools to:

- Search for information in a library containing assigned readings, supporting resources provided by an instructor, and artifacts the learner creates (e.g., terms, notes).
- Select content in an assigned reading to highlight it or tag it.
- Make a note structured by a schema that records the annotation in a web form with slots tailored to a schema – e.g., TERM NOTE: term, definition, example, see also ...; or DEBATE NOTE: claim, evidence, warrant, counterclaim, my position.
- Organize items in a folder-like directory.

Phase 4, strategic revision of tactics and strategies for learning, is not included in Table 21.2; it is addressed in the later section on Learning Analytics for SRL.

As Winne and Baker (2013) noted, "Self-regulated learning (SRL) is a behavioural expression of metacognitively guided motivation" (p. 3). Consequently, every trace records a motivated choice about how to learn. Beyond representing features of the COPES model, traces reveal learners' beliefs about worthwhile effort that operationalizes choices among alternative goals.

The Learner's Reports

Paper-based questionnaires (surveys) and live oral reports are prevalent choices of methods for gathering data about learning events. Oral reports can be obtained through interviews outside the temporal boundaries of a studying session or during learning-on-the-fly as think aloud reports.

In both paper-based (or electronically presented) questionnaires and oral reports, learners are prompted to describe one or more features of COPES. The nature of the prompt is critical because it establishes several external conditions that a co-operative learner uses to set standards for deciding what to report. A thorough review is beyond the scope of this chapter; see Winne and Perry (2000) and Winne (2010b). In general, because questionnaire data are only weakly contextual (e.g., When you study, how often do you/how important is it for you to ...?) and because all forms of self-report data suffer loss, distortion, and bias due to frailties of human memory, they may not reliably indicate how a learner goes about learning in any particular study episode or how learning varies (is self-regulated) as conditions vary. Self-report data are important, however, because they do reliably reflect beliefs learners hold about COPES. Beliefs shape what learners attend to about tasks, about themselves, and about standards they set for themselves.

Materials Studied

Materials learners work with are sources of data about conditions that may bear on how they engage in SRL. Texts can be described by various analytics including

Table 21.2. Illustrative Traces and Inferences about Phases of SRL

Phase of SRL	Trace	Inference
1) Survey resources and constraints	Search for "marking rubric" or "requirements" at the outset of a study episode.	An internal condition, namely, a learner's expectation that guidance is available about the requirements for a task.
	Open several documents, scan each for 15–30 s, close.	Refreshing information about previous work, if documents were previously studied; or scanning for particular but unknown information.
2) Plan and set goals	Start timer.	Plan to metacognitively monitor pace of work.
	Fill in fields of a "goal" note with slots: goal, milestones, indicators of success.	Assemble a plan in which goals are divided into sub-goals (milestones), set standards for metacognitively monitoring progress.
3) Engagement	Select and highlight content.	Metacognitive monitoring, unknown standards.
	Select and tag content.	Metacognitive monitoring; the standard used to monitor is revealed by the tag applied (e.g., confusing, good point).
	Select a bigram (e.g., greenhouse gas, slapstick comedy) and create a term.	Metacognitive monitoring content for technical terminology, assembling the term with a definition.
	Select content and annotate it using a "debate note" form, filling in slots: claim, evidence, warrant, counterclaim, my position.	Metacognitive monitoring with the standard to test whether content is an argument + assemble and rehearse information about the argument.
	Open a note created previously.	Metacognitive monitoring knowledge relative to a standard of completeness or accuracy, judge knowledge does not meet the standard.
	Put documents and various notes into a folder titled "Project Intro."	Metacognitively monitor uses of content; The standard is "useful for the introduction to a project"; assembling elements in a plan for future work.

readability[2] and cohesion (e.g., Coh-Metrix[3]). Content can be indexed for the extent to which learners have had opportunity to learn it plus characteristics of what a learner learned from previous exposures. Materials a learner studies also can be identified for the presence of rhetorical features such as examples and multichannel presentations of information, such as a quadratic expression described in words (semantic), an equation (symbolic), and a graph (visual) forms.

LEARNING ANALYTICS FOR SRL

Learning analytics to support SRL typically will have two elements: a calculation and a recommendation. The calculation – e.g., notation about presence, count, proportion, duration, probability – is based on traces of actions carried out during one or multiple study episodes (Roll & Winne, 2015a). A numeric report may be conveyed along with or as a visualization. Examples might be a stacked bar chart showing relative proportions of highlights, tags and notes created while studying each of several web pages, a timeline marked with dots that show when particular traces were generated, and a node-link graph depicting relations among terms in a glossary (link nodes when one term is defined using another term) with heat map decorations showing how often each term was operated on while studying. This element directly or by transformation mirrors information describing COPES traced in the history of a learner's engagement. Table 21.3 presents illustrative trace data that might be mirrored.

2 See, for example, http://www.wordscount.info/readability.html
3 http://cohmetrix.com/

A "simple" history of trace data mirrored back to a learner may be conditioned or contextualized by other data: features of materials such as length or a readability index, demographic data describing the learner (e.g., prior achievement, hours of extracurricular work, postal code), or other characterizations of learners such as disposition to procrastinate, degree in a social network (the number of people with whom this learner has exchanged information) or context for study (MOOC vs. face-to-face course delivery, opportunity to submit drafts for review by peers before handing in a final copy to be marked).

The second element of a learning analytic about SRL is a recommendation – what should change about how learning is carried out plus guidance about how to go about changing it. Learners can directly control three facets of COPES: operations, standards, and some conditions (Winne, 2014). Products are controllable only indirectly because their characteristics are function of 1) conditions a learner is able to and chooses to vary, particularly information selected for operations; and, 2) which operation(s) the learner chooses to apply in manipulating information. Evaluations are determined by the match of product attributes and the particular standards a learner adopts for those products. Recommendations about changing conditions, operations, or standards may be grounded in findings from data mining not guided by theory, by findings from research

Table 21.3. Analytics Describing COPES Facets in SRL

COPES	Description
Conditions	Presence/absence of a condition within a learning episode Onset/offset along the timeline in a study episode or across a series of episodes
Operations	Frequency of SMART operations (see Table 21.1) Sequence, pattern, conditional probability one SMART operation relative to others
Product	Presence Completeness (e.g., number of fields with text entered in a note's schema) Quality
Standard	Presence Precision Appropriateness
Evaluation	Presence Validity

in learning science, nor a by combination.

Whether a recommendation is offered or not, change in the learner's behaviour traces the learner's evaluation that 1) previous approaches to learning were not sufficiently effective or satisfactory and 2) the learner predicts benefit by adopting the recommendation or an adaptation of it. In this sense, learning analytics update prior external conditions and afford new internal conditions. Together, a potential for action is created, but this is only a potential for two reasons. First, learners may not know how or have skill to enact a recommendation. Second, because learners are agents, *they* control their learning. As Winne and Baker (2013) noted:

> What marks SRL from other forms of regulation and complex information processing is that the goal a learner seeks has two integrally linked facets. One facet is to optimize achievement. The second facet is to optimize how achievement is constructed. This involves navigating paths through a space with dimensions that range over processes of learning and choices about types of information on which those processes operate. (p. 3)

Thus, learning analytics afford opportunities for learners to exercise SRL but the learner decides what to do. There is an important corollary to this logic. If a learning analytic is presented without a recommendation for action, an opportunity arises for investigating options a learner was previously able to exercise on his or her own and, now, chooses to exercise. In other words, motivation and existing tactics for learning can be assessed by analytics that omit recommendations and guidance for action.

CHALLENGES FACING LEARNING ANALYTICS ABOUT SRL

Research on learning analytics as support for SRL is nascent. The field has just begun to map frontiers, including what to trace, instrumentation for gathering traces, interfaces that optimize gathering data without overly perturbing learning activities, computational tools for constructing analytics about SRL that meld trace data with other data, scheduling delivery of learning analytics, and features of information conveyed in learning analytics (Baker & Winne, 2013; Roll & Winne, 2015b). Amidst these many topics, several merit focused exploration.

Grain Size

Features of learning events can be tracked at multiple grain sizes ranging from individual keystrokes and clicks executed along a timeline marked off in very fine time units (e.g., tens of milliseconds) to quite coarse grain sizes (e.g., the URL of a web page and when it loads, the learner's overall score on a multi-item practice quiz). Different methods for aggregating fine-grained data will represent features of COPES differently. While this affords multiple views of how learners engage in SRL, several questions arise.

First, how will depictions of SRL and recommendations for adapting learning vary across learning analytics formed from data at different grain sizes? An analogy might be made to chemistry. Chemical properties and models of chemical interactions vary depending on whether the unit is an element, a compound, or a mixture. Consider two grain sizes for information that is manipulated with an assembling operation: 1) snippets of text selected for tagging when studying a web page, and 2) entire artifacts – quotes, notes, and bookmarks – that a learner files in a titled (tagged) folder. Future research may reveal that assembling at one grain size has different implications for learning relative to assembling at another grain size.

If grain size matters, one implication is that approaches to forming learning analytics may benefit by considering not only whether and which operations are applied – what a learner does – but also characteristics of information to which operations are applied.

Learning analytics for SRL may benefit by blending counts and other quantitative descriptions of COPES with semantic, syntactic, and rhetorical features of conditions, products, and standards.

Because coarser grained reflections of SRL generally, but not necessarily, are built up using finer grained data, another issue arises in developing and using statistical calculations. Statistical descriptions that describe relationships among larger-grained features of learning and SRL, such as correlation and distance metrics, may share finer-grained constituents. This inherently introduces part-whole relationships. Will that matter?

Time

Excepting research in learning science that investigates how achievement covaries with time spans between episodes of studying, reviewing, and taking tests (Delaney et al., 2010), the phenomenon of forgetting (Murayama, Miyatsu, Buchli, & Storm, 2014) and loss of knowledge across the summer vacation (Cooper, Nye, Charlton, Lindsay, & Greathouse, 1996), time data has been underused. Traces and other data available to learning analytics commonly can be supplemented with time stamps. Much research remains to investigate how temporal features of COPES and coarser-grained descriptors may play useful roles in learning analytics about SRL as a process that unfolds within each studying episode and across a series of episodes. One focus for this research is identifying patterns in COPES events across time (Winne, Gupta, & Nesbit, 1994). Vexing questions here are how to define the span of a time window within which patterns are sought and the degree to which non-focal events intervening in an encompassing pattern can be identified and filtered out (see Zhou, Xu, Nesbit, & Winne, 2011). Another key topic relating to time is investigating when learning analytics should be delivered: in real time (i.e., approximately instantaneously following an event or identification of a pattern), on demand (by learners or instructors), or at punctuated intervals (e.g., weekly)?

Generalization

Learning science strives to balance the accuracy of descriptions about particular learning events in contrast to describing how learning events relate to outcomes, which requires ignoring details to allow generalizing over specific events. When data and time stamps at very fine-grain sizes are available about the course of studying over time, accuracy of description is maximized. How should generalizations be formed, tested, and validly interpreted as accuracy is deliberately compromised (see Winne, 2017)?

The goal of education is development – of knowledge, interest, confidence, critical thinking, and so on. If education succeeds, each learner changes over time, and changes quite likely vary among peers. Even if there is genuinely big data, at very fine grain sizes of data, it is statistically very unlikely any two learners' data signatures perfectly match. Learning analytics face a challenge to find balance between accuracy and generalization when describing one learner's ipsative development or the match of that learner's "learning signature" to others'. The field of learning analytics will benefit from frequent consideration of this challenge.

ACKNOWLEDGEMENTS

This work was supported by grants to Philip H. Winne from the Canada Research Chairs Program, and the Social Sciences and Humanities Research Council of Canada 410-2011-0727 and 435 2016-0379.

REFERENCES

Arnold, K. E., & Pistilli, M. D. (2012). Course Signals at Purdue. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (LAK '12), 29 April–2 May 2012, Vancouver, BC, Canada (pp. 267–270). New York: ACM. doi:10.1145/2330601.2330666

Azevedo, R., Moos, D. C., Johnson, A. M., & Chauncey, A. D. (2010). Measuring cognitive and metacognitive regulatory processes during hypermedia learning: Issues and challenges. *Educational Psychologist*, 45(4), 210–223.

Baker, R. S. J. d., & Winne, P. H. (Eds.). (2013). Educational data mining on motivation, metacognition, and self-regulated learning [Special issue]. *Journal of Educational Data Mining*, 5(1).

Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research*, 66(3), 227–268.

Delaney, P. F., Verkoeijen, P. P., & Spirgel, A. (2010). Spacing and testing effects: A deeply critical, lengthy, and at times discursive review of the literature. *Psychology of Learning and Motivation*, 53, 63–147.

Dunlosky, J., & Rawson, K. A. (2015). Practice tests, spaced practice, and successive relearning: Tips for classroom use and for guiding students' learning. *Scholarship of Teaching and Learning in Psychology*, 1(1), 72–78.

Educause (n.d.) Next generation learning initiative. http://nextgenlearning.org

Eckerson, W. W. (2006). *Performance dashboards: Measuring, monitoring, and managing your business*. Hoboken, NJ: John Wiley & Sons.

Elias, T. (2011, January). *Learning analytics: Definitions, processes and potential*. http://learninganalytics.net/LearningAnalyticsDefinitionsProcessesPotential.pdf

Murayama, K., Miyatsu, T., Buchli, D., & Storm, B. C. (2014). Forgetting as a consequence of retrieval: A meta-analytic review of retrieval-induced forgetting. *Psychological Bulletin*, 140(5), 1383–1409.

Roll, I., & Winne, P. H. (2015a). Understanding, evaluating, and supporting self-regulated learning using learning analytics. *Journal of Learning Analytics*, 2(1), 7–12.

Roll, I., & Winne, P. H. (Eds.). (2015b). Self-regulated learning and learning analytics [Special issue]. *Journal of Learning Analytics*, 2(1).

Siemens, G. (2010, August 25). *What are learning analytics?* http://www.elearnspace.org/blog/2010/08/25/what-are-learning-analytics/

Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (1966). *Unobtrusive measures*. Skokie, IL: Rand-McNally.

Winne, P. H. (1995). Inherent details in self-regulated learning. *Educational Psychologist*, 30, 173–187.

Winne, P. H. (1997). Experimenting to bootstrap self-regulated learning. *Journal of Educational Psychology*, 89, 397–410.

Winne, P. H. (2010a). Bootstrapping learner's self-regulated learning. *Psychological Test and Assessment Modeling*, 52, 472–490.

Winne, P. H. (2010b). Improving measurements of self-regulated learning. *Educational Psychologist*, 45, 267–276.

Winne, P. H. (2011). A cognitive and metacognitive analysis of self-regulated learning. In B. J. Zimmerman and D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 15–32). New York: Routledge.

Winne, P. H. (2014). Issues in researching self-regulated learning as patterns of events. *Metacognition Learning*, 9(2), 229–237.

Winne, P. H. (2017). Leveraging big data to help each learner upgrade learning and accelerate learning science. *Teachers College Record*, 119(3). http://www.tcrecord.org/Content.asp?ContentId=21769.

Winne, P. H. (in press). Cognition and metacognition in self-regulated learning. In D. Schunk & J. Greene (Eds.), *Handbook of self-regulation of learning and performance*, 2nd ed. New York: Routledge.

Winne, P. H, & Baker, R. S. J. d. (2013). The potentials of educational data mining for researching metacognition, motivation and self-regulated learning. *Journal of Educational Data Mining*, 5(1), 1–8.

Winne, P. H., Gupta, L., & Nesbit, J. C. (1994). Exploring individual differences in studying strategies using graph theoretic statistics. *Alberta Journal of Educational Research*, 40, 177–193.

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277–304). Mahwah, NJ: Lawrence Erlbaum Associates.

Winne, P. H., & Perry, N. E. (2000). Measuring self-regulated learning. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 531–566). Orlando, FL: Academic Press.

Zhou, M., Xu, Y., Nesbit, J. C., & Winne, P. H. (2011). Sequential pattern analysis of learning logs: Methodology and applications. In C. Romero, S. Ventura, S. R. Viola, M. Pechenizkiy & R. Baker (Eds.), *Handbook of educational data mining* (pp. 107–121). Boca Raton, FL: CRC Press.