Computers & Education 104 (2017) 18–33

Contents lists available at ScienceDirect

Computers & Education

journal homepage: www.elsevier.com/locate/compedu

Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses

René F. Kizilcec [a, *, 1], Mar Pérez-Sanagustín [b, 2], Jorge J. Maldonado [b, c, 2]

[a] Department of Communication, Stanford University, USA
[b] Department of Computer Science, Pontificia Universidad Católica de Chile, Chile
[c] Department of Computer Science, Universidad de Cuenca, Ecuador

ARTICLE INFO

Article history:
Received 9 June 2016
Received in revised form 31 August 2016
Accepted 1 October 2016
Available online 4 October 2016

Keywords:
Online learning
Learning analytics
Individual differences
Self-regulated learning
Massive open online course

ABSTRACT

Individuals with strong self-regulated learning (SRL) skills, characterized by the ability to plan, manage and control their learning process, can learn faster and outperform those with weaker SRL skills. SRL is critical in learning environments that provide low levels of support and guidance, as is commonly the case in Massive Open Online Courses (MOOCs). Learners can be trained to engage in SRL and actively supported with prompts and activities. However, effective implementation of learner support systems in MOOCs requires an understanding of which SRL strategies are most effective and how these strategies manifest in online behavior. Moreover, identifying learner characteristics that are predictive of weaker SRL skills can advance efforts to provide targeted support without obtrusive survey instruments. We investigated SRL in a sample of 4,831 learners across six MOOCs based on individual records of overall course achievement, interactions with course content, and survey responses. We found that goal setting and strategic planning predicted attainment of personal course goals, while help seeking was associated with lower goal attainment. Learners with stronger SRL skills were more likely to revisit previously studied course materials, especially course assessments. Several learner characteristics, including demographics and motivation, predicted learners' SRL skills. We discuss implications for theory and the development of learning environments that provide adaptive support.

1. Introduction

A primary goal of Massive Open Online Courses (MOOCs) is to provide more people with opportunities for personal and intellectual growth. Between late 2011 and 2015, 550 institutions created 4,200 courses that reached over 35 million people worldwide, according to data collected by Class Central (Shah, 2015). Most learners who enroll in MOOCs selectively engage with parts of the course content and a small proportion eventually completes the course (Anderson, Huttenlocher, Kleinberg, & Leskovec, 2014; Breslow et al., 2013; Evans, Baker, & Dee, 2016; Ho et al., 2015; Kizilcec, Piech, & Schneider, 2013; Perna et al., 2014; Seaton, Bergner, Chuang, Mitros, & Pritchard, 2014). This variation in behavior can be partly attributed to the

* Corresponding author.
E-mail addresses: kizilcec@stanford.edu (R.F. Kizilcec), mar.perez@ing.puc.cl (M. Pérez-Sanagustín), jjmaldonado@uc.cl (J.J. Maldonado).
[1] Present address: Department of Communication, 450 Serra Mall, Stanford University, Stanford, CA 94305, USA.
[2] Present address: Department of Computer Science, Pontificia Universidad Católica de Chile, Avda. Vicuña Mackenna 4860, Macul, Santiago, Chile.

http://dx.doi.org/10.1016/j.compedu.2016.10.001
0360-1315/

remarkable diversity of leaners' backgrounds, motivations, intentions, and prior experiences (de Barba, Kennedy, & Ainley, 2016; Kizilcec & Schneider, 2015; Littlejohn, Hood, Milligan, & Mustain, 2016; Zheng, Rosson, Shih, & Carroll, 2015). In fact, only half of the survey respondents in a typical MOOC report that they intend to complete the course to receive a certificate (Kizilcec & Schneider, 2015; Littlejohn & Milligan, 2015; Reich, 2014). However, even among learners who have ambitious goals for the course and who are committed to achieve them, a majority of learners is unsuccessful. The primary reasons for attrition in MOOCs are related to poor time management and course difficulty, according to both quantitative (Kizilcec & Halawa, 2015; Nawrot & Doucet, 2014) and qualitative (Zheng, Rosson, Shih, & Carroll, 2015) research.

In the absence of support and guidance from an instructor, the ability to regulate one's learning process is a critical skill to achieve personal learning objectives. Unlike in school settings, where time is typically structured around classes and everyone follows a fixed schedule, online learners need to determine when and how to engage with course content of their own accord. Prior work found that many learners struggle with self-regulation in online learning environments (Lajoie & Azevedo, 2006). In the context of MOOCs, which afford low levels of support and guidance, the absence of external pressure to make progress and explicit social norms around completion requires that learners be highly self-directed to achieve their course goals (Banerjee & Duflo, 2014; Hew & Cheung, 2014; Kizilcec & Halawa, 2015; Zheng et al., 2015). This raises the question of how to support learners to achieve their goals in learning environments like MOOCs.

To address this question, we investigated self-regulation strategies in MOOCs. Our work builds on self-regulated learning (SRL) theory, which describes ways for learners to take control of their learning process. We examined which self-regulation strategies predict attainment of personal course goals, how different strategies manifest in records of interactions with course content, and how strategies vary by individual characteristics. The goal of this work is to provide a foundation for future research and interventions that support SRL in MOOCs and comparable environments. We used MOOCs as an environment in which to investigate authentic learner behavior over time—a research paradigm that holds promise for advancing educational science and practice (Reich, 2015; Winne & Nesbit, 2010)—in combination with methods from educational data mining and learning analytics (Roll & Winne, 2015; Winne & Baker, 2013). We surveyed 4,831 online learners across six distinct MOOCs about their SRL strategies and individual characteristics, (e.g., demographics, motivations, and intentions for completing course materials). Their responses were combined with detailed records of their interactions with course content and their overall course achievement, yielding a longitudinal account of SRL in an authentic learning context.

This article makes two contributions to the literature on SRL. First, we provide insight into SRL and its behavioral manifestations in MOOCs for a heterogeneous adult learner population. Second, leveraging the heterogeneity of the present sample, we demonstrate multiple individual differences in SRL that can inform targeted interventions, such as adaptive scaffolding. In Section 2, to develop our research questions, we review the literature on models of SRL and research in online learning contexts, including MOOCs. The methodology of this study is described in Section 3. Results are presented in Section 4. Section 5 offers a summary of the main findings, their implications and limitations. We conclude with reflections on future research.

2. Related work

This section provides a review of relevant literature on SRL. First, we review different models of SRL and justify our choice of Pintrich's (2000) model for this investigation. Then, we review literature on SRL in online environments to identify relevant SRL strategies and narrow the scope of our investigation to the most effective strategies. Finally, we develop three research questions that we address empirically.

2.1. Models of self-regulated learning

The large body of literature on SRL, which has developed over the last two decades, encompasses numerous definitions of SRL (Puustinen & Pulkkinen, 2001) and models to explain SRL (Boekaerts, 1999; Borkowski, 1996; Butler & Winne, 1995; Pintrich, 2000; Winne & Hadwin, 1998; Zimmerman, 2000). Self-regulated learners are characterized by their ability to initiate metacognitive, cognitive, affective, motivational, and behavioral processes in order to take actions to achieve their learning goals and persevere until they succeed. Despite variation in terminology, most scholars assume SRL "to proceed from some kind of a preparatory or preliminary phase, through the actual performance or task completion phase, to an appraisal or adaptation phase" (Puustinen & Pulkkinen, 2001, p. 280). Moreover, researchers agree that SRL is not a fixed trait, but rather a skill that can be developed and honed through experience and practice applying SRL strategies (Azevedo & Cromley, 2004; Schunk, 2005; Zimmerman, 2015). Notably, feedback is an inherent catalyst for self-regulated behavior. Learners generate internal feedback and process external feedback, for instance, by setting criteria for success and monitoring their engagement with tasks relative to these criteria (Butler & Winne, 1995).

Two established models of SRL stand out in the literature as distinct approaches to explain the same process: Pintrich's model focuses on different kinds of SRL strategies, while Zimmerman's model disentangles SRL into three phases: forethought, performance, and self-reflection.[3] Pintrich (2000, p. 453) defines SRL as "an active, constructive process whereby

[3] Winne and Hadwin (1998) proposed a 4-phase model of SRL predicated on the view that learners are agents who exert choice over their learning process. This model essentially separates the forethought phase into two, one in which learners scan their environment for relevant resources and constraints, and another in which they set goals.

learners set goals for their learning and then attempt to monitor, regulate and control their cognition, intentions and behavior, guided and constrained by their goals and the contextual features of the environment." In contrast, Zimmerman (2000, p. 14) describes SRL as "self-generated thoughts, feelings, and actions that are planned and cyclically adapted to the attainment of personal goals." Both models have been influential in the literature and share common ideas. Here we follow in the tradition of Pintrich's model, because its focus on particular strategies lends itself more to large-scale quantitative investigations that can inform targeted interventions to support specific SRL strategies.

Predicated on the notion that self-regulation can be trained, Pintrich identified three categories of SRL strategies that learners can apply to regulate their learning: (1) cognitive, (2) metacognitive, and (3) resource management strategies. Learners utilize cognitive strategies in the acquisition, storage, and retrieval of information (e.g., rehearsal, critical thinking, organization, elaboration). Learners utilize metacognitive strategies to plan, monitor and regulate their learning process to accomplish a goal (e.g., goal setting and strategic planning, self-monitoring, and self-evaluation). Learners utilize resource management strategies to manage the learning environment and external resources (e.g., time management, help seeking, effort regulation and organizing one's study environment).

2.2. *Self-regulated learning in online learning and MOOCs*

In this section, we review literature to identify the most effective SRL strategies in online learning environments. Compared to in-person instruction, contemporary online learning environments tend to provide learners with less support and guidance on how to learn deeply yet efficiently. Online learners are expected to actively and autonomously engage in the learning process (C.-H. Wang, Shannon, & Ross, 2013), which demands both a high level of confidence in their own abilities and the ability to manage their own learning process (Liang & Tsai, 2008; Sun & Rueda, 2012; Tsai, Chuang, Liang, & Tsai, 2011). Learners who struggle to regulate their learning process effectively tend to experience frustration and become less engaged in the course (Sun & Rueda, 2012), and they are ultimately less successful (Lee, Shen, & Tsai, 2008; Samruayruen, Enriquez, Natakuatoong, & Samruayruen, 2013; Tsai, 2009).

Prior work in online learning environments demonstrated improvements in academic achievement from applying SRL strategies, especially time management, metacognition, and effort regulation strategies (Azevedo & Aleven, 2013; Broadbent & Poon, 2015; Niemi, Nevgi, & Virtanen, 2003). These strategies help learners process and retain knowledge in a structured manner (Beishuizen & Steffens, 2011; Dignath & Büttner, 2008; Pintrich, 2004; Zimmerman, 2008). Several studies found that providing scaffolding for these strategies can support SRL and raise achievement (Azevedo, Moos, Greene, Winters, & Cromley, 2008; Kim & Hodges, 2012; Taub, Azevedo, Bouchet, & Khosravifar, 2014). Supportive feedback, such as clarifying to learners what constitutes good performance and encouraging positive motivational beliefs, can support SRL and it can be implemented in online learning environments at large scale (Nicol & Macfarlane-dick, 2006).

The recent availability of large and fine-grained datasets has led to investigations at the intersection of SRL and learning analytics. The assessment of frequencies and sequences of regulatory activities in learning environments provides a novel perspective on SRL that complements and potentially supersedes traditional self-report measures (Bannert, Reimann, & Sonnenberg, 2014; Beheshitha, Gašević, & Hatala, 2015; Hadwin, Nesbit, Jamieson-Noel, Code, & Winne, 2007; Winne, 2014). For example, Siadaty, Gašević, and Hatala (2016) investigated how scaffolding interventions via a recommender system can support SRL in a workplace and found informative discrepancies between self-reported and actual behavior. Process mining techniques can offer new insight into how learners self-regulate in MOOCs, for instance by examining differences between designed versus observed learning paths (Davis, Chen, Hauff, & Houben, 2016).

Studies of attrition in MOOCs suggest that metacognitive strategies and resource and task management strategies are critical for success. In particular, learners' reasons for disengaging from MOOCs can inform which SRL strategies are potentially valuable. Kizilcec & Halawa (2015) examined reasons for disengaging in a sample of 1698 learners across 20 MOOCs and identified four clusters (in order of significance): time issues, course difficulty, format and content, and goals and expectations. A follow-up study, targeted at learners predicted to have recently disengaged from a MOOC, confirmed this pattern using open-ended responses that were coded: 84% of respondents mentioned that they had "not enough time for the course." Half of the 84% who faced time issues also indicated being easily distracted from the course, which suggests that better metacognitive and resource management strategies could have prevented their disengagement. Additionally, satisfaction and relative progress in the MOOC were associated with goal striving, a critical antecedent to goal setting and strategic planning. In another study, Zheng et al. (2015) interviewed 18 learners about their experience in MOOCs and the issue of not having enough time also emerged in their analysis. Moreover, the lack of pressure in this non-traditional environment emerged as another factor influencing persistence, which relates to task management strategies, such as effort regulation. Learners also missed a sense of community, which would limit the use of help-seeking strategies. Finally, Nawrot and Doucet (2014) found further evidence that the most common reasons for disengaging from MOOCs can be related to task management (e.g., time management) and metacognitive strategies (e.g., strategic planning, goal setting).

Besides reasons for disengaging from MOOCs, what insights can successful learners offer about strategies that were helpful? Interviews of 17 learners who successfully completed a MOOC helped identify several ostensibly effective strategies (Kizilcec, Pérez-Sanagustín, & Maldonado, 2016). A number of them were task management strategies, such as reserving time in the week for studying (time management), starting and finishing a chapter on the same day (task strategies, effort regulation), and working with others on the course (help seeking). Other reported strategies were metacognitive strategies, such as having clear objectives and planning around those (goal setting, strategic planning), applying what one has learned in the course to internalize it, and creating summaries or mind maps of lecture content (self-evaluation, self-monitoring). This account of learner-generated strategies complements the findings on reasons for attrition.

Overall, based on the findings of prior work on self-regulation in online learning environments, and MOOCs in particular, we focus on metacognitive strategies and resource and task management strategies. Specifically, we consider the following strategies that are expected to support learners in MOOCs (exemplary definitions from the literature are provided):

1 **Goal setting**: Setting of educational goals or sub-goals in order to exert the effort required to achieve those goals (Schunk, 2005; Zimmerman, 2000).
2 **Strategic planning**: Planning the sequence, timing, and completion of activities directed at learning goals (Zimmerman & Pons, 1986).
3 **Self-evaluation**: Setting quality standards and criteria for progress to judge one's own performance (Boud, 1995). Activities for monitoring the learning process in relation to defined learning goals (Schunk, 2005).
4 **Task strategy**: Organizing, planning, and transforming one's own study time (time management) and tasks (i.e., timing, sequencing, pacing, rearrangement of instructional materials) (Effeney, Carroll, & Bahr, 2013; Zimmerman & Pons, 1986). Activities to improve persistence and effort-regulation in the face of academic challenge (Richardson, Abraham, & Bond, 2012).
5 **Elaboration**: Combining new knowledge with prior knowledge and constructing meaning from learned materials (Niemi et al., 2003). Extending or modifying the learning materials to make them more meaningful and memorable (Weinstein, Acee, & Jung, 2011).
6 **Help seeking**: Asking other people for help, such as the instructor or one's peers, or consulting external help and resources (Pintrich, 1999; Richardson et al., 2012).

2.3. Individual differences in self-regulated learning and online course behavior

Prior research has investigated how individual differences between learners might relate to both self-reported SRL and behavior in MOOCs. First, in terms of self-reported SRL, learners who report higher levels of motivation, commitment to learn, formal education, and relevant prior knowledge also indicate higher levels of SRL (Hood, Littlejohn, & Milligan, 2015; Littlejohn et al., 2016). Hood et al. (2015) examined how learners' context (i.e., background characteristics) influences their ability to self-regulate their learning in MOOCs. They found higher levels of SRL among learners with a higher level of formal education and among working professionals in domains related to the course content. Littlejohn et al. (2016) found differences between learners with varying levels of SRL in their reported motivations and goals for the course, which apparently shaped their approach to the MOOC and their use of learning strategies. On the basis of in-depth interviews, they identified differences in self-described learning behaviors between learners with low versus high SRL profiles for five SRL sub-processes. By contrast, several investigations have found no significant gender differences in terms of SRL in the context of various digital learning environments (Basol & Balgalmis, 2016; Liou & Kuo, 2014; Yukselturk & Top, 2013).

Second, numerous studies have found individual differences in learners' engagement and achievement in MOOCs. Empirical investigations have linked variation in course behavior and achievement with various individual differences: learners' demographic and personal background (Evans et al., 2016; Guo & Reinecke, 2014; Hansen & Reich, 2015; Kizilcec & Halawa, 2015), motivations for enrolling and intentions for the course (de Barba et al., 2016; Jordan, 2014; Kizilcec & Halawa, 2015; Kizilcec & Schneider, 2015; Reich, 2014), and self-efficacy (Wang & Baker, 2015). Guo and Reinecke (2014) analyzed the navigation strategies of course certificate earners by age and country of origin. They found older learners and learners from countries with fewer teachers per student to take less linear paths through the course content, which could be a sign of lower SRL skills. Based on a sample of over 67,000 learners across 16 MOOC, Kizilcec and Halawa (2015) found higher grades and levels of persistence among male learners, and those with more formal education, stronger time commitment to the course, prior experience with the course topic, an intent to complete the course, and who were located in the Global North. Across 68 courses, Hansen and Reich (2015) found that U.S. learners with lower socioeconomic resources were also less likely to enroll in and complete MOOCs, especially among adolescents and young adults. To summarize, prior work has identified individual differences in terms of SRL and in terms of behavior and achievement in MOOCs. Thus, in a context with a highly heterogeneous learner population, individual differences warrant further empirical investigation.

2.4. Research questions

The current literature offers several accounts of SRL in MOOCs and individual differences based on characteristics such as learners' formal education, prior knowledge, and their professional context. This prior work provides a basis for deeper investigations of SRL in large-scale online learning environments. We identified two gaps in our current understanding of SRL in online learning that warrant further investigation.

The first gap concerns our understanding of the relation between self-reported SRL strategies and objective behavioral measures in a large-scale learning environment over time. How does learning behavior differ between highly self-regulated learners and less self-regulated ones? Prior work suggests that learners' self-reported SRL strategies influence how they behave in MOOCs, but most prior studies only examined SRL in small-scale online environments. Yet how SRL manifests in learners' actual interactions with course content in MOOCs has received no scholarly attention. Moreover, we found no evidence on the relative efficacy of different SRL strategies to support online learners achieve personal learning goals over time. We identified six SRL strategies that have been related to academic achievement in online learning and MOOCs in prior work (see Section 2.2). However, the relative extent to which these SRL strategies predict differences in achieving personal goals in MOOCs is unknown. We therefore pose the following two research questions:

RQ1. Which self-reported SRL strategies are most helpful to achieve personal course goals?
RQ2. How do self-reported SRL strategies manifest in interactions with course content?

The second gap in the literature concerns our understanding of individual differences in SRL. Prior work found individual characteristics of learners such as their level of education, gender, age, course intentions, and motivations to be associated with performance in the course. For example, prior investigations have demonstrated that learners with more formal education self-report stronger SRL skills and exhibit higher persistence and achievement (Hood et al., 2015: Kizilcec & Halawa, 2015). However, there has not been a systematic analysis of individual characteristics that predict learners' self-reported SRL, because this demands a large and diverse survey sample of learners, which is rarely available outside of MOOCs. Insight into individual differences in SRL can support efforts to develop targeted intervention, for example, by using this information to set Bayesian priors in models. We will identify a broad set of individual differences in SRL in terms of characteristics, many of which were examined in prior work (demographics, course intentions, motivations, etc.) to investigate the following research question:

- **RQ3.** How do self-reported SRL strategies vary by individual learner characteristics?

3. Methods

3.1. Participants and context

The final study sample included 4,831 online learners in six distinct MOOCs on topics in Engineering, Computer Science, Management, Transportation, and Education. The courses were offered by Pontificia Universidad Católica de Chile through Coursera were taught in Spanish and followed a self-paced format, such that course materials were available all at once without deadlines. Each course encompassed 6–10 sections, each containing 5–10 video lectures and several assessments (e.g., multiple-choice quizzes, peer-review activities). Most course assessments were formative and could be attempted multiple times. The target audiences of these courses were high school & college students and professionals in subject-related industries. To improve the generalizability of our findings, we selected courses with a range of topics that would in turn attract a diverse learner audience. Indeed, based on self-reported demographics, the average age was 32.0 (SD = 10.8), 26% were women, 63% held a bachelor's or higher degree (15% a master's or Ph.D.), 60% were employed, and 25% were students. Data was collected between April and December 2015.

The final sample is a subset of the 6,709 learners who answered the initial course survey about their SRL strategies and various individual characteristics, including demographics, course intentions and motivations. The following exclusions were made: First, 385 responses were removed from learners who either took the survey more than once in the same course, or completed virtually no survey questions. Another 1,450 responses had to be excluded because it was not possible to combine the survey data with corresponding course data (data entry errors resulted in unmatched responses, because we relied on information that was manually entered by survey respondents to merge the datasets). The remaining 4,874 responses came from 4,831 unique learners, as 43 learners were enrolled in two of the six courses. To reduce complexity in the analysis, we randomly selected one of the courses for each of these 43 learners.

3.2. Measures

Participants completed an optional course survey when entering the course for the first time. The survey asked learners about their demographics (age, gender, education, occupation), time commitment (hours per week), course intentions (intend to watch all lectures; intend to complete all assessments), prior experience with the course topic, the number of prior online courses started, and the number of completed courses. The survey also included the Online Learning Enrollment Intentions (OLEI) scale (Kizilcec & Schneider, 2015) translated into Spanish.[4] Finally, the survey included a measure of SRL that was adapted from the questionnaires used by Littlejohn and Milligan (2015) and Barnard, Paton, and Lan (2008), which are based on several established instruments (Barnard-Brak, Paton, & Lan, 2010; Pintrich & others, 1991; Rigotti, Schyns, & Mohr, 2008; Schraw & Dennison, 1994; Warr & Downing, 2000). Based on our review of SRL strategies in online learning environments (see Section 2.2), we selected six strategy subscales from the original instrument (items previously used by Azevedo

[4] Spanish translation of the OLEI provided at http://dx.doi.org/10.6084/m9.figshare.1585144.

et al., 2008; Taub et al., 2014). The resulting questionnaire (see Appendix) had participants rate 23 statements about SRL strategies on how characteristic they were for them on a labeled 5-point scale (coded 0 to 4): goal setting strategies (4 statements), strategic planning (4), self-evaluation (3), task strategies (6), elaboration (3), and help seeking (3). The order in which statements were presented in the survey was randomized.

The individual score for each strategy was computed by averaging ratings of corresponding items. Table 1 provides descriptive statistics for the collected SRL survey data with an exemplary statement for each strategy and a composite computed by averaging scores for all strategies. The SRL measure had high reliability for all strategy subscales with Cronbach's α of at least 0.75, despite the small number of items used. As shown in Table 1, the help-seeking subscale had a lower mean and lower correlation with the composite; this may be partly because it was the only subscale that included a reverse-coded item. A small amount of missing responses (fewer than 5%) was imputed using predictive mean matching (Little, 1988), a method that accounts for the joint distribution of observations on all other items to predict the most likely value for a missing observation; in general, this technique yields lower non-response bias than simple mean imputation or discarding incomplete responses (cf. Buuren & Groothuis-Oudshoorn, 2011).

3.3. Analytic approach

We addressed **RQ1** about the **relationship between SRL and achieving personal course goals** by assessing associations between each strategy and course outcomes depending on learners' stated course goal. We used non-parametric Spearman correlation coefficients, because the outcome data was either binary or skewed. Additionally, we fitted logistic regression models to evaluate the predictive power of the six SRL strategies simultaneously.

To investigate **RQ2** about **how SRL strategies manifest in interactions with course content**, we computed relevant variables that characterize learners' interactions in the course. First, we preprocessed records of learners' interactions with course materials to identify individual sessions, defined here as sustained periods of activity during which a learner interacts with the course materials at least once in two hours. Then, within each individual session, consecutive interactions with the same object in the course materials were aggregated and labeled based on learners' progress with the object. Table 2 provides definitions for the different interaction states. Interactions with any course object lasting for less than five seconds were excluded to reduce noise. The resulting dataset was then aggregated in two ways for analysis: first at the level of individual transitions, we computed the frequency of transitions from one interaction state to another within sessions, and second at the level of sessions, we computed activity metrics during each session in terms of (i) time spent on different content types or progress states, (ii) the number of materials interacted with during the session, and (iii) the time spent between sessions.

Transition-level analysis. For the analysis of individual transitions between interaction states, as defined in Table 2, we counted the frequency of direct state transitions, yielding a six by six transition matrix for each learner. The transition counts were transformed into transition probabilities and normalized for each initial interaction state. Fig. 1 illustrates this transition graph. For clarity, only arrows for transition probabilities greater than 15% are shown and the transition probability is indicated next to each arrow. We evaluated correlations between learners' reported SRL strategies and their individual transition probabilities for the 36 possible transitions. This yielded a total of $6 * 36 = 216$ Spearman correlation coefficients. To focus on the most robust patterns, we only considered highly significant correlations ($p < 0.01$, based on 10,000 course-clustered bootstrap replications). Of the 34 highly significant correlations, 14 correlations were relatively large in magnitude (i.e., $|r| > 0.045$; max. $|r| = 0.065$).

Session-level analysis. For the analysis of per-session activity, we computed several quantities for each session, including time spent overall and specifically on i. course assessments, ii. revisiting any content, iii. revisiting assessments, and iv. revisiting lectures. Then, aggregating over sessions, we considered (a) the total time as the sum of the above session-based times, (b) the proportion of total time spent on specific activities out of the overall total time spent, and (c) the median time spent on each activity to characterize a typical session. Conceptually, these three summary statistics—sum total, proportion, and median—characterize learner engagement in terms of how much time was spent overall, how a given amount of time was spent, and how much was typically spent, respectively. Additionally, we considered the total number of sessions; the median number of lectures and assessments that learners engaged with; and finally, the median and SD of durations between subsequent sessions. As most of these variables are undefined for a single observation, this analysis only included learners who had more than one recorded session ($N = 2{,}949$). We followed the same analytic approach as in the transition-level analysis, but this time we evaluated 6 strategies * 22 behavioral indicators = 132 correlation coefficients. Bootstrapping yielded 27 highly significant correlations, 18 of which were of relatively large magnitude (i.e., $|r| > 0.045$; max. $|r| = 0.097$).

To address **RQ3** about **individual differences in self-reported SRL strategies**, we considered 27 individual learner characteristics. The self-reported characteristics encompassed learners' demographics (8 predictors) and time commitment, their experience with the course topic, their prior experience with online courses (2 predictors), and their goals for the course (2 predictors) and motivations for enrolling (13 predictors). We used penalized regression to identify individual characteristics that were most predictive of each SRL strategy. The advantage of penalized regression in this context is that it performs variable selection. The algorithm shrinks coefficients on predictor variables that provide little or no improvement to model fit, thereby effectively excluding unimportant predictors from the model. Another advantage of this approach over evaluating individual differences separately for each variable is that the regression coefficients are estimated simultaneously. When considering individual characteristics that are correlated, such as age and education, the estimated coefficients characterize the predictor's association with an SRL strategy while adjusting for all other predictors in the model. Continuous predictors (age, online courses started/finished, time commitment) were standardized to zero mean and unit variance. All remaining predictors were binary and dummy-coded for the analysis. Scores for the six SRL strategy outcomes were also standardized for ease of interpretation. We applied an elastic net penalty (Zou & Hastie, 2005) in the regression models, which performs variable selection akin to the LASSO penalty (Tibshirani, 1996), but it is less prone to randomly choosing between highly correlated predictors. We used a 90% LASSO with 10% Ridge penalty and 10-fold cross-validation to identify the parameter value that minimized the prediction mean-squared error (cf. Friedman, Hastie, & Tibshirani, 2001). The penalized regression models yielded six sets of coefficients that are illustrated in Fig. 2.

4. Results

We begin with general observations about the survey results. Learners reported an average time commitment of 4.9 hours per week ($SD = 3.1$; median $= 4$). The vast majority reported an intention to watch all lectures (95%) and complete all assessments (93%) in the course. Half of the learners reported having prior experience with the course topic and a majority had prior experience with online courses (number of prior online courses started: $M = 2.4$, $SD = 4.0$, median $= 1$; number of completed courses: $M = 1.8$, $SD = 3.2$, median $= 1$). The most pronounced SRL strategies reported were self-evaluation and elaboration, followed by strategic planning, task strategies, and goal setting; the least common strategy was help seeking (Table 1). Moreover, several of the SRL strategies were highly correlated, such as goal setting with strategic planning ($r = 0.70$), strategic planning with task strategies ($r = 0.66$), and task strategies with elaboration ($r = 0.72$). Help seeking was the least correlated strategy with the overall SRL composite.

4.1. Which self-reported SRL strategies are most helpful to achieve personal course goals? (RQ1)

We evaluated how SRL strategies were related to achieving three different personal course goals: first, earning a course certificate, which required achieving satisfactory grades on course assessments; second, completing assessments (independent of grades), and third, watching lectures in the course. For each personal goal, we assessed the correlation between self-reported SRL strategies and goal attainment among those who expressed the goal. Results are provided in Table 3. We found that goal setting and strategic planning were significant positive predictors of goal attainment for all three goals. In contrast, help seeking was a significant negative predictor of goal attainment (except for completing lectures, $p = 0.069$). Self-evaluation and task strategies were predictive only of completing assessments and lectures, while elaboration was not at all correlated with goal attainment.

In light of high correlations between strategies, we proceeded to fit logistic regression models to evaluate all six SRL strategies simultaneously when predicting goal attainment. Goal setting was a strong positive predictor of goal attainment, while help seeking was a strong negative predictor. Results were consistent across personal course goals and robust to regression adjustment for available covariates (demographics, experience, commitment, etc.), and notably, strategic planning was also a strong positive predictor with goal setting excluded from the model. For example, learners who indicated 1 SD higher levels of goal setting had 54% higher odds of achieving their goal of earning a certificate ($z = 2.68$, $p = 0.007$). By contrast, the same model yielded 27% lower odds of certification ($z = -3.11$, $p = 0.002$) for learners who indicated 1 SD higher levels of help seeking. Likewise, coefficient estimates predicting the other course goals were highly significant and only somewhat smaller. Thus, learners who engaged in goal setting and avoided help seeking were significantly more likely to achieve their personal course goals. Although several other SRL strategies were individually associated with goal attainment, goal setting and help seeking emerged as the two key predictors.

Table 1
Descriptive statistics for each SRL strategy and an average SRL composite ($\bar{\mathbf{x}}$) with exemplary statements, mean and standard deviation, Chronbach's α, and pairwise Pearson's correlation coefficients.

Strategy	Example statement	M (SD)	α	2.	3.	4.	5.	6.	$\bar{x}$
1. Goal Setting	I set realistic deadlines for learning.	3.0 (0.76)	0.86	0.70	0.48	0.57	0.46	0.29	0.78
2. Strategic Planning	I organize my study time to accomplish my goals to the best of my ability.	3.1 (0.65)	0.75		0.60	0.66	0.58	0.32	0.84
3. Self-evaluation	I think about what I have learned after I finish.	3.3 (0.66)	0.80			0.63	0.59	0.25	0.74
4. Task Strategies	When I study for this course, I make notes to help me organize my thoughts.	3.1 (0.62)	0.78				0.72	0.35	0.87
5. Elaboration	When I am learning, I try to relate new information I find to what I already know.	3.3 (0.64)	0.77					0.32	0.77
6. Help Seeking	When I do not understand something, I ask others for help.	2.6 (0.79)	0.77						0.58
$\bar{x}$ SRL Composite	–	3.0 (0.52)	0.92						

Table 2
Definitions of interaction states with course materials to characterize consecutive learner behavior.

Interaction State	Definition
Lecture Begin	Begin but not complete watching a lecture that was not previously completed.
Lecture Complete	Complete watching a lecture for the first time. The lecture may have been begun before but not completed.
Lecture Revisit	Watch (part of) a lecture that was completely watched in the past.
Assessment Attempt	Attempt but not pass an assessment that was not previously passed.
Assessment Pass	Pass an assessment for the first time. The assessment may have been attempted before but not passed.
Assessment Revisit	Revisit (re-take or simply look up) an assessment that was already passed.

Table 3
Associations between achieving personal course goals and SRL strategies in terms of Spearman correlation coefficients evaluated for binary certification outcome and continuous proportion of assessments/lectures completed in the course.

Personal course goal	Expressed goal (and attained goal)	Goal setting	Strategic planning	Self-evaluation	Task strategies	Elaboration	Help seeking
Earn course certificate	32% (8.9%)	0.08**	0.05*	≈0	0.04	0.03	−0.05*
Complete all assessments	93% (7.3%)[a]	0.05**	0.05**	0.04*	0.04*	0.03	−0.05**
Complete all lectures	95% (9.1%)[a]	0.03*	0.04**	0.03*	0.03*	0.03	−0.03

[a] Goal attainment was evaluated for completing over 80% of assessments and lectures, respectively. *$p < 0.05$; **$p < 0.005$.

4.2. How do self-reported SRL strategies manifest in interactions with course content? (RQ2)

Manifestations of self-reported SRL were evaluated first at the level of individual transitions and then based on per-session activity. The two analytic approaches are meant to provide complementary perspectives on how SRL manifests in course behavior.

First, we considered learner behavior at the level of individual transitions between interaction states, as defined in Table 2. The transition graph in Fig. 1 shows that, after completing a lecture, learners were most likely to complete another lecture (in 74% of cases) or pass an assessment (13%; arrow omitted in Fig. 1). In comparison, learners who just passed an assessment were most likely to complete a lecture next (66%). Those who attempted an assessment without passing it went on to either attempt (29%) or pass (19%) a different assessment, or complete (21%) or revisit (27%) a lecture. The remaining transitions depicted in Fig. 1 can be interpreted the same way. To assess how differences in SRL manifest in individual transitions, we evaluated correlations between SRL strategies and individual transition probabilities for the six strategies and 36 possible transitions (out of the 6 * 36 = 216 correlations, 14 stood out as significant and relatively strong patterns; see Section 3.5). Results are summarized in Table 4 with descriptions of the 14 identified transition patterns by SRL strategy. Overall, learners who report stronger SRL skills were more inclined to revisit course materials after completing other materials, instead of starting new materials. An exception to this overarching trend: learners inclined to seek help were less likely to pass an assessment after completing a lecture.

Second, we analyzed course behavior based on per-session activity. In comparison to the analysis of individual transitions, the per-session activity analysis provided coarser but more contextualized behavioral indicators, including detailed information about the total amount, the proportion of time, and the median amount of time spent on various course activities across sessions. Note that the distinction between total time, proportion of time, and median time is meaningful:

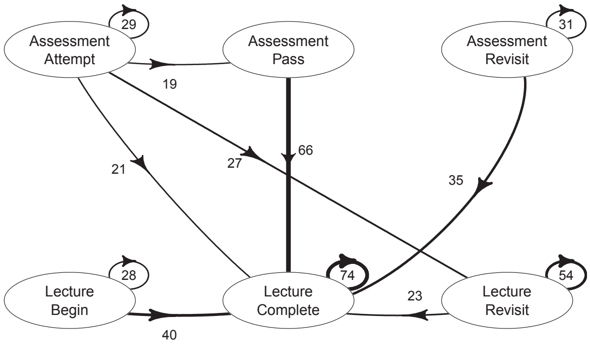

Fig. 1. Transition graph of interactions with course content for all learners. Interaction states are defined by content type and progress level (cf. Table 2). Arrow thickness is proportionate to transition probability, which is provided as percentages next to the arrowhead. Only showing arrows for probabilities above 15%.

Table 4
Manifestation of SRL strategies at the level of individual transitions and per-session activity. Upward and downward arrows indicate significant positive and negative correlations, respectively.

SRL strategy	Manifestations at the level of individual transitions	Manifestation at the level of per-session activity
Goal Setting	↑ revisiting an assessment after passing an assessment or completing a lecture ↑ revisiting a lecture after completing a lecture	↑ total time spent and number of sessions ↑ time spent – total and proportional – on assessments and revisiting assessments ↑ median number of assessments taken and median time spent on assessments
Strategic Planning	↑ revisiting an assessment or lecture after passing an assessment	↑ time spent – total and proportional – on revisiting assessments ↑ median time spent on assessments
Self-evaluation	↑ revisiting an assessment after passing an assessment or completing a lecture ↑ revisiting a lecture after beginning a lecture ↑ beginning a lecture after revisiting a lecture	↑ time spent – total and proportional – on revisiting assessments
Task Strategies	↑ revisiting an assessment or lecture after passing an assessment ↑ beginning a lecture after revisiting a lecture	↑ total time spent revisiting content ↑ proportion of time spent revisiting assessments ↑ median time spent on revisiting assessments
Elaboration	↑ revisiting an assessment after completing a lecture	No significant correlations
Help Seeking	↓ passing an assessment after completing a lecture	No significant correlations

while the total amount reveals how much time someone spends on specific materials, the proportion of time reveals how someone spends their time in the course; for instance, devoting relatively more time to assessments than lectures is expected to enhance learning (Koedinger, Mclaughlin, Kim, Zhuxin Jia, & Bier, 2015). The median complements the other two metrics by characterizing activity in a typical session. We evaluated correlations between six SRL strategies and 22 behavioral indicators (out of the $6 * 22 = 132$ correlations, 18 stood out as strong and significant patterns; see Section 3.5). Results are summarized in Table 4 with descriptions of the 18 identified patterns by SRL strategy. Overall, learners who reported using more goal setting, strategic planning, self-evaluation, and task strategies spent more time in absolute and relative terms on revisiting assessments. Moreover, learners who reported using more goal setting and strategic planning also spent more time on assessments in a typical session; learners who reported using more task strategies devoted more time to revisiting lectures.

4.3. How do self-reported SRL strategies vary by individual learner characteristics? (RQ3)

We assessed individual differences in self-reported SRL strategies based on 27 individual characteristics, encompassing demographics, prior experience, time commitment, goals and motivations. Fig. 2 illustrates the results of six penalized regressions, one for each SRL strategy, with coefficient estimates from each model in each column. Blank entries in Fig. 2 indicate that the penalized regression estimation shrunk a coefficient to zero, thereby excluding the corresponding predictor from the model. Estimates are adjusted for all other predictors in the model; for example, the coefficient on age is estimated adjusting for all other characteristics in the model, such as occupation and level of education.

A number of individual differences emerged for learner demographics. Older learners reported consistently higher levels of SRL, except for help seeking. Women reported lower levels of strategic planning, elaboration, and self-evaluation; however, women reported higher levels of goal setting, task strategies, and especially help seeking. Compared to the 37% of learners in the sample who had not earned at least a bachelor's degree, those with a bachelor's degree reported lower strategic planning, self-evaluation, and help seeking. By contrast, learners with a professional or master's degree, and especially those with a Ph.D. reported higher levels of goal setting, strategic planning, and task strategies. While learners with a Ph.D. reported generally strong SRL skills, they reported being much less inclined to seek help. Learners who were also students in school or university reported consistently lower SRL, especially for self-evaluation and task strategies. In contrast, learners who were employed were more inclined to engage in goal setting, strategic planning, and help seeking, despite lower levels of self-evaluation.

Individual differences by learners' prior experience were more consistent across strategies. Learners who had started more online courses in the past consistently reported lower SRL, while those who had completed more online courses consistently reported higher SRL, especially goal setting. Those with prior experience with the course topic reported higher levels for most SRL strategies, but were less inclined to seek help. Furthermore, learners who were willing to commit more time to the course reported consistently higher SRL. Likewise, SRL skills were substantially higher—up to 0.5 SD—among learners who expressed the goal of either finishing all lectures or finishing all assessments.

Finally, SRL also varied by learners' motivations for enrolling in the course. The following enrollment intentions were generally associated with higher SRL skills: enrolling to earn a certificate, to meet new people, to take the course with others, for the prestige of the institution or instructor, or because the course is relevant to one's research, one's job, or one's school/

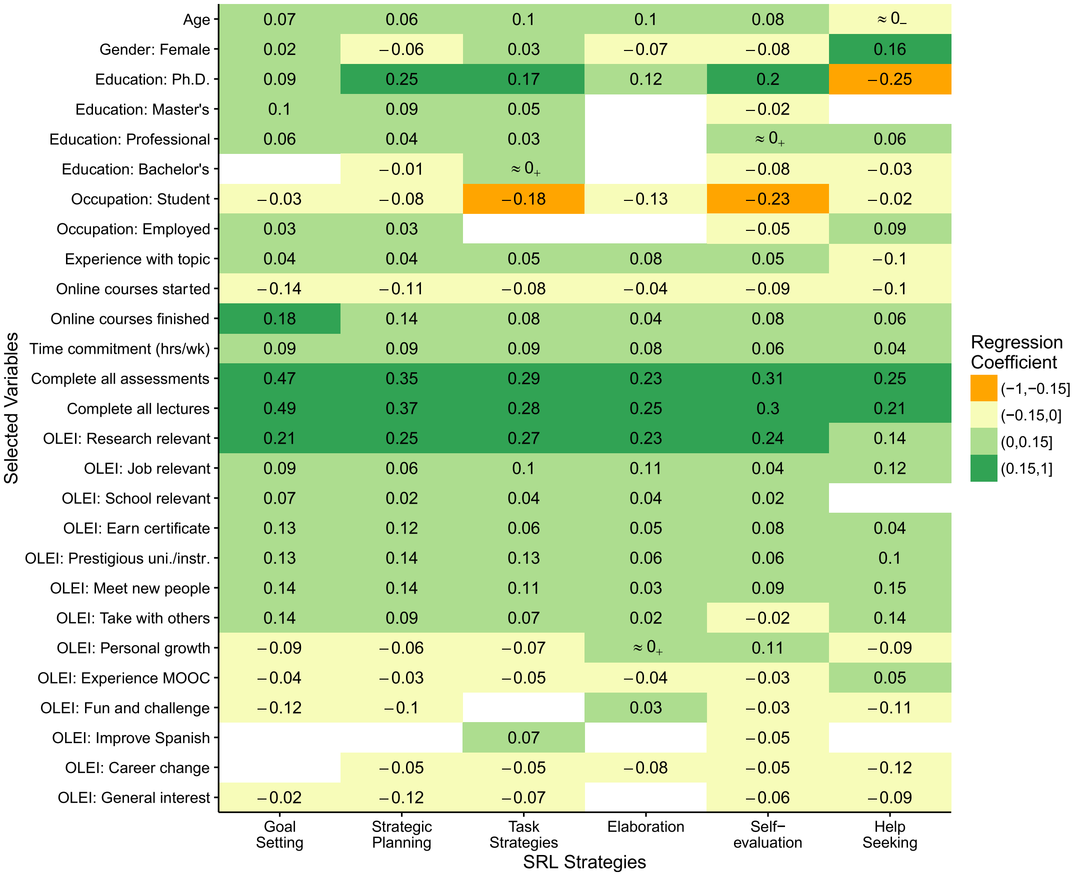

Fig. 2. Individual differences in SRL examined by demographics, prior experience, time commitment, goals and motivations (marked OLEI). Showing penalized regression coefficients for six models, one for each SRL strategy, with standardized continuous predictors (i.e., age, online courses started/finished, time commitment) and dummy-coded binary predictors (all other predictors). SRL outcome variables were also standardized for ease of interpretation. Blank boxes indicate predictor variables that were excluded by variable selection. Colors indicate the sign and magnitude of coefficients.

degree program. By comparison, the following enrollment intentions were generally associated with lower SRL skills: enrolling out of general interest, for career change, for fun and challenge, to experience a MOOC, or for personal growth.

5. Discussion

This study provides a quantitative account of SRL that advances our understanding of which SRL strategies support online learners in MOOCs, how different strategies manifest behaviorally in the learning environment, and how SRL strategies vary across a heterogeneous group of learners. Our results are based on an analysis of survey and platform log data from 4,831 learners across six MOOCs. We briefly summarize the findings pertaining to each of the three research questions that we investigated. First, which self-reported SRL strategies are most helpful to achieve personal course goals? (RQ1) We found that learners who reported engaging more in goal setting and strategic planning were more likely to attain personal course goals, such as earning a certificate—consistent with prior research on these strategies (e.g., Schunk, 2005; Zimmerman & Pons, 1986; Zimmerman, 2000). In contrast, help seeking was a negative predictor of goal attainment, unlike in prior work (Pintrich, 1999; Richardson et al., 2012). Interpretations of these findings are discussed below. Second, how do self-reported SRL strategies manifest in interactions with course content? (RQ2) Drawing on two operationalizations of interactions with course content (individual transitions and per-session activity), we found that high levels of self-reported SRL (except for help seeking) manifested in frequent revisiting of course materials—especially course assessments—that were previously completed. However, those inclined to seek help were actually less likely to pass assessments after lectures. Third, how do self-reported SRL strategies vary by individual learner characteristics? (RQ3) A large number of significant individual differences in self-reported SRL were found. Gender differences emerged in the use of multiple SRL strategies; in particular, women were more inclined to seek help than men, in contrast to prior work that found gender differences (Basol & Balgalmis, 2016; Liou & Kuo, 2014; Yukselturk & Top, 2013). Learners with a Ph.D. were generally more self-regulated, but much less inclined to seek help. In contrast, learners who were also students reported lower SRL, especially for self-evaluation and task strategies. Learners with ambitious course intentions, greater time commitment, and prior experience with the topic generally indicated stronger SRL skills. Finally, motivations for taking the course that signaled a relevant and supportive life context (taking course with a friend, course relevant to job/school/research, etc.) predicted stronger SRL skills, while motivations that signaled a less supportive context (taking course for fun and challenge, to experience a MOOC, for career change, etc.) predicted weaker SRL skills.

5.1. Implications

The present findings have implications for theory and practice around SRL in the context of MOOCs and similar online learning environments. We discuss four implications of our findings in the context of prior work: (1) supporting goal setting and strategic planning; (2) interpreting behavioral manifestations of SRL; (3) interpreting the negative results for help seeking, and (4) leveraging insights from individual differences.

First, goal setting and strategic planning stood out as particularly helpful strategies in MOOCs. Learners who reportedly engaged in these metacognitive strategies were more likely to achieve their course goals and engaged more deeply with course assessments, perhaps because they also appreciate the value of assessments for checking their understanding and receiving feedback to support their learning. The results are consistent with accounts from prior work that highlight goal setting and strategic planning as important factors underlying attrition and achievement in MOOCs (Kizilcec & Halawa, 2015; Kizilcec, Piech, & Schneider, 2016; Nawrot & Doucet, 2014; Zheng et al., 2015). According to the analysis of individual differences, older learners with more formal education who would be expected to have more developed metacognitive abilities indeed reported engaging more in goal setting and strategic planning. In light of this converging correlational evidence that goal setting and strategic planning support learners, MOOCs should provide learners with relevant scaffolding to support these strategies at the beginning of the course and throughout as needed. The vast literature on ways to support goal pursuit offers many types of interventions, some of which with established efficacy in academic settings. For example, mental contrasting with implementation intentions (MCII) is a metacognitive self-regulation strategy that is known to promote academic performance in children and self-discipline in adults, at least in the USA and Germany (Duckworth, Grant, Loew, Oettingen, & Gollwitzer, 2011; Duckworth, Kirby, Gollwitzer, & Oettingen, 2013). MCII can be implemented as a brief set of writing activities that prompt reflection about positive outcomes and obstacles related to a goal, followed by making if-then plans about how to handle obstacles in the future. Guiding learners through activities such as MCII holds promise for supporting learners in MOOCs, especially those with weaker metacognitive skills. The implementation of this type of scaffolding is associated with negligible costs, and it can scale to any number of learners. Moreover, it provides an opportunity to study how to best support goal pursuit with a highly heterogeneous population.

Second, our analysis of behavioral manifestations of SRL strategies revealed multiple behavioral indicators. The finding that learners with strong SRL skills (except for help seeking) tended to revisit course content, especially assessments, suggests that learners engaged in retrieval practice—repeatedly testing their knowledge to promote learning (Karpicke & Roediger, 2008). Although the behavioral indicators were correlated with specific self-reported SRL strategies, they did not necessarily reflect applications of the strategies themselves. Intuitively, it is reasonable that learners who report engaging in self-evaluation also spent more time revisiting assessments; however, the interpretation is not always clear and the activity log provides only an indirect account of when and how learners apply specific SRL strategies. Based on our bottom-up approach to discover behavioral manifestations (see e.g., Hadwin et al., 2007), a learner could have engaged in multiple different SRL strategies, given the observation that they revisited course assessments. By contrast, a top-down approach would first categorize (sequences of) actions indicative of specific strategies and then observe these instances of strategies in the data (see e.g., Beheshitha et al., 2015). Each approach has its merits. Our bottom-up approach to the discovery of behavioral manifestations of SRL was in-keeping with the exploratory nature of this work, as it reduced the extent to which findings were constrained by researcher expectations. The results can inform the design of future confirmatory studies that adopt a top-down approach.

Third, the finding that help seeking negatively predicts goal attainment can be interpreted several ways. It seems surprising considering that prior work has found that learners who report working on the course with someone else, such as a friend, have higher performance (e.g., Breslow et al., 2013; Kizilcec & Schneider, 2015). Although the two constructs were positively related in our analysis of individual differences (see Fig. 2, 'OLEI: Take with others'), which suggests some degree of overlap, they differ in several ways. Learners who have coordinated with a 'study buddy' are probably very organized and committed to the course, engage in collaborative learning, and benefit from mutual support and social accountability. Learners who reported being more inclined to seek help were perhaps alone in their educational endeavor and were hoping to enter an active community of learners who support each other during the course. While the courses in this study technically had discussion forums, they were either not used at all or featured fewer than a dozen posts—thus, there was no online community of learners, at least not in the course environment. This environment therefore did not provide adequate support for learners with a preference for collaborative learning and who expect to be able to ask for help when they face challenges. Alternatively, learners may report a high inclination for help seeking, because they are less confident in their ability to succeed in the course (i.e., low self-efficacy; Bandura, 1997), a belief that may be justified given that there are no pre-requisites required for taking a MOOC. The observed negative association between help seeking and goal attainment could be either due to self-selection or a sign that learners' help-seeking needs were not met in the learning environment. Additionally, the finding that help seeking negatively predicts completing assessments after completing lectures suggests that learners who are not inclined to seek help may be more confident in their abilities. As a result, they are more motivated to test their knowledge with assessments and complete them successfully. Prior work on intelligent tutoring systems, where learners can request hints from the system at any time, found that help seeking does not necessarily support learning, because learners may not request hints at the right time (Aleven, Stahl, Schworm, Fischer, & Wallace, 2003). Moreover, norms around help seeking vary across cultural contexts, with more organic helping behavior occurring in more collectivist contexts (Ogan et al., 2015). Thus, effective scaffolding for help seeking in MOOCs may need to adapt to different contexts, and in addition to conveying the potential benefits of help seeking, it should provide guidance on when to choose productive struggle over seeking help.

Fourth, our findings of individual differences in SRL between learners who expressed different motivations for taking the course provide empirical evidence consistent with recent work. Hood et al. (2015) also found increased self-reported self-regulated learning behaviors among learners who studied or worked in a field related to the course topic compared to those without a topic-relevant role or context. Littlejohn et al. (2016) conducted in-depth interviews with MOOC learners and found consistent evidence for the role of learners' context in shaping their perceptions of their learning process and the purpose of the course. Specifically, learners with a relevant professional context reported higher SRL skills and expressed motivations related to professional development, which allowed them to view MOOCs as a non-formal learning opportunity. The present study confirms these findings with a large set of contextual indicators derived from learners' enrollment intentions. Our findings suggest that individual differences in terms of commonly available learner characteristics (e.g., course intentions, education level, employment status, gender) could be leveraged in combination with real-time behavioral data to provide adaptive scaffolding. The design of scaffolding certainly needs to be informed by causal evidence to determine which nudges and embedded support systems actually improve learning outcomes and course goal attainment (for examples of SRL interventions in MOOCs that did not have a positive impact, see Davis, Chen, Zee, Hauff, & Houben, 2016; Kizilcec et al., 2016). The scale and level of heterogeneity in MOOCs offers an unprecedented opportunity to advance our understanding of what are the critical dimensions for adaptive SRL scaffolding. Thanks to increasingly large and diverse learner samples, minority subpopulations that would historically be aggregated with a majority group can receive the scholarly attention that they deserve, without compromising scientific rigor (Kizilcec & Brooks, in press).

5.2. Limitations

The present study has three notable limitations that, although common in this type of work, should be noted when drawing conclusions from our findings. The first limitation concerns external validity. The study is based on a sample of mostly Latin American learners engaging in MOOCs that were offered in Spanish. On the one hand, given that most published findings are based on samples from Western educated industrialized rich democratic countries (Henrich, Heine, & Norenzayan, 2010), this study advances the inclusivity of our science by drawing on a non-traditional sample. On the other hand, as noted here and in prior work, learners' socio-cultural context has consequences for how they perceive and engage with online courses (e.g., Guo & Reinecke, 2014; Ogan et al., 2015). While our findings are consistent with prior work that considered other international populations, future work should replicate and extend the current findings with other samples to test generalizability. The same argument applies to the specific courses that were studied, which were self-paced MOOCs on the Coursera platform in 2015. Prior work found differences between the staggered versus all-at-once content release format for MOOCs in terms of persistence and completion in the course (Mullaney & Reich, 2015). While the courses covered a wide range of topics, the design and instrumentation of the platform at the time are expected to play an important role in shaping learner behavior and researchers' interpretation of their behavior through the lens of the collected data. This highlights a structural limitation with implications for both the replicability of findings across platforms and time, and the reliability of inferences that can be drawn from meta-analyses of related research findings.

The second limitation concerns construct validity. The instrument we used to assess SRL is based on established and validated instruments in the literature. However, we did not employ any one complete instrument. Instead, we identified six relevant SRL strategies from prior literature and adapted established instruments to specifically measure the selected constructs. This approach made a trade-off between utilizing a complete instrument with many items that are unsuitable in the MOOC context, on the one hand, and creating entirely new survey items to measure established constructs, on the other hand. It is simply unreasonable to ask an online volunteer learner population to fill out a lengthy battery of survey questions and expect to receive data that is of high quality. Another consideration regarding measurement is that we translated the entire survey into Spanish, including the measure of SRL. The translation was performed by two native speakers who understood the underlying constructs that were assessed. Valid translation of survey instruments is a non-trivial issue and it warrants empirical validation.

A third limitation is that most instantiations of SRL strategies considered in this study could not be observed directly in the MOOC environment—no data was available about whether learners set clear learning goals, engaged in note-taking while watching lectures, practiced self-explanation, or consulted friends or the Internet for help. Unless SRL strategies are facilitated in the environment or through linked third-party applications, neither self-report nor course log data provides a complete account of online SRL and therefore limits the ability to draw valid conclusions about SRL in these environments. Recent work has combined MOOC data with information from relevant online platforms (e.g., StackExchange, GitHub, and LinkedIn) to find evidence of learning having an impact outside of the course (Chen, Davis, Lin, Hauff, & Houben, 2016). While these online platforms are not instrumented to assess SRL either, they may offer insight into how learners' skills progress over time during, after, and in-between online courses.

6. Conclusion

This study contributes to laying the groundwork for an educational scientific evolution of leveraging new data sources and methodologies to advance educational theory and practice. Until recently, most educational researchers had the choice of investigating small-scale but rich learner data (e.g., classroom observation, laboratory studies) or large-scale but shallow learner data (e.g., standardized test scores, field surveys). The research opportunities in digital learning environments promise to deliver the best of both worlds, rich and large-scale educational data with the ability to learn from randomized experiments at a rapid pace (cf. Reich, 2015). An important challenge in this domain is finding effective ways to connect empirical work back to theory to contribute to generalizable knowledge. This concern has been raised by leading scholars in learning analytics in general and specifically in the context of data-intensive research on SRL (Gašević, Dawson, & Siemens, 2015; Winne, 2014). Observational accounts of SRL in *in vivo* online learning environments, such as the current study, offer a window into the reality of SRL over time for online learners from diverse backgrounds. This type of research marks the necessary transition from established theory and the empirical work that supports it to applications in novel contexts that offer different affordances and concern broader populations. The resulting correlational findings permit refinement of theory and can inform the next wave of experimental research (e.g., identifying SRL interventions that help learners in MOOCs) and subsequent evaluations of adaptive systems (e.g., identifying heterogeneous treatment effects to understand which intervention works when for whom). To this end, our research findings offer an account of SRL in MOOCs that quantifies the relative benefits of different learning strategies, identifies behavioral manifestations of different strategies, and exposes individual differences in self-reported SRL.

Finally, we highlight three promising directions for future research in this area. First, the development of predictive models of when learners fail to apply SRL strategies that would support their learning—this can inform small but timely interventions (e.g., prompting a learner with "Have you tried explaining this concept to yourself?"). A second direction is the development of feedback systems that facilitate self-monitoring of SRL strategies for learners with weak metacognitive skills. This will require new approaches to reliably quantify SRL engagement from behavioral traces. A third research direction concerns interventions to support SRL for a global and diverse learner population. It may be necessary to adapt strategies for different social and cultural contexts to be effective. Research to address these three challenges will serve to advance our current understanding of SRL in online environments and inform new strategies to better support learners.

Acknowledgements

This work was supported by FONDECYT (Chile) [grant number N 11150231], the MOOC-Maker Project [grant number 561533-EPP-1-2015-1-ES-EPPKA2-CBHE-JP], and the Comisión Nacional de Investigación Científica - CONICYT Ministry of Education, Chile, Ph.D. Student Fellowships and University of Cuenca, Ecuador. We thank Dan Davis and two anonymous reviewers for their comments on the manuscript. Researchers who would like to obtain the dataset used in this study for further investigation can request it from the corresponding author.

Appendix. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.compedu.2016.10.001.

References

Aleven, V., Stahl, E., Schworm, S., Fischer, F., & Wallace, R. (2003). Help seeking and help design in interactive learning environments. *Review of Educational Research, 73*(3), 277–320.

Anderson, A., Huttenlocher, D., Kleinberg, J., & Leskovec, J. (2014). Engaging with massive online courses. *WWW '14 Proceedings of the 23rd International Conference on World Wide Web*, 687–698. http://dx.doi.org/10.1145/2566486.2568042.

Azevedo, R., & Aleven, V. (2013). Metacognition and learning technologies: An overview of current interdisciplinary research. In *In International handbook of metacognition and learning technologies* (pp. 1–16). New York: Springer.

Azevedo, R., & Cromley, J. G. (2004). Does training on self-regulated learning facilitate students' learning with hypermedia? *Journal of Educational Psychology, 96*(3), 523.

Azevedo, R., Moos, D. C., Greene, J. A., Winters, F. I., & Cromley, J. G. (2008). Why is externally-facilitated regulated learning more effective than self-regulated learning with hypermedia? *Educational Technology Research and Development, 56*(1), 45–72.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY, USA: Freeman.

Banerjee, A. V., & Duflo, E. (2014). (Dis)Organization and success in an economics MOOC. In *American Economic Review* (Vol. 104, pp. 514–518). http://dx.doi.org/10.1257/aer.104.5.514.

Bannert, M., Reimann, P., & Sonnenberg, C. (2014). Process mining techniques for analysing patterns and strategies in students' self-regulated learning. *Metacognition and Learning, 9*(2), 161–185. http://dx.doi.org/10.1007/s11409-013-9107-6.

de Barba, P. G., Kennedy, G. E., & Ainley, M. D. (2016). The role of students' motivation and participation in predicting performance in a MOOC Motivation and participation in MOOCs. *Journal of Computer Assisted Learning*, (March 2015)http://dx.doi.org/10.1111/jcal.12130.

Barnard-Brak, L., Paton, V. O., & Lan, W. Y. (2010). Profiles in self-regulated learning in the online learning environment. *The International Review of Research in Open and Distributed Learning, 11*(1), 61–80.
Barnard, L., Paton, V., & Lan, W. (2008). Online self-regulatory learning behaviors as a mediator in the relationship between online course perceptions with achievement. *The International Review of Research in Open and Distributed Learning, 9*(2).
Basol, G., & Balgalmis, E. (2016). A multivariate investigation of gender differences in the number of online tests received-checking for perceived self-regulation. *Computers in Human Behavior, 58*, 388–397.
Beheshitha, S. S., Gašević, D., & Hatala, M. (2015). A process mining approach to linking the study of aptitude and event facets of self-regulated learning. *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge - LAK '15*, 265–269. http://dx.doi.org/10.1145/2723576.2723628.
Beishuizen, J., & Steffens, K. (2011). A conceptual framework for research on self-regulated learning. In *Self-regulated learning in technology enhanced learning environments* (pp. 3–19). Springer.
Boekaerts, M. (1999). Self-regulated learning : Where we are today. *International Journal of Educational Research, 31*, 445–457.
Borkowski, J. G. (1996). Metacognition: Theory or chapter heading? *Learning and Individual Differences, 8*(4), 391–402.
Boud, D. (1995). *Enhancing learning through self-assessment*. Routledge.
Breslow, L., Pritchard, D. E., DeBoer, J., Stump, G., Ho, A. D., & Seaton, D. (2013). Studying learning in the worldwide classroom: Research into edX's first MOOC. *Research & Practice in Assessment, 8*(March 2012), 13–25. Retrieved from http://www.mendeley.com/catalog/studying-learning-worldwide-classroom-research-edxs-first-mooc/.
Broadbent, J., & Poon, W. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *The Internet and Higher Education, 27*, 1–13. http://dx.doi.org/10.1016/j.iheduc.2015.04.007.
Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research, 65*(3), 245–281. http://dx.doi.org/10.3102/00346543065003245.
Buuren, S., & Groothuis-Oudshoorn, K. (2011). Mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software, 45*(3).
Chen, G., Davis, D., Lin, J., Hauff, C., & Houben, G.-J. (2016). Beyond the MOOC platform: Gaining Insights aboutLearners from the Social Web. *Proceedings of the 8th ACM Conference on Web Science, WebSci'16*, 15–24. http://dx.doi.org/10.1145/2908131.2908145.
Davis, D., Chen, G., Hauff, C., & Houben, G. (2016). Gauging MOOC Learners' Adherence to the Designed Learning Path. In *Proceedings of the 9th International Conference on Educational Data Mining (EDM)*. Raleigh, NC, USA.
Davis, D., Chen, G., Zee, T., Van Der, Hauff, C., & Houben, G. (2016). Retrieval Practice and Study Planning in MOOCs: Exploring Classroom-Based Self-Regulated Learning Strategies at Scale. In *Proceedings of the 11th European Conference on Technology Enhanced Learning (EC-TEL)*.
Dignath, C., & Büttner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning, 3*(3), 231–264.
Duckworth, A. L., Grant, H., Loew, B., Oettingen, G., & Gollwitzer, P. M. (2011). Self-regulation strategies improve self-discipline in adolescents: Benefits of mental contrasting and implementation intentions. *Educational Psychology, 31*(1), 17–26. http://dx.doi.org/10.1080/01443410.2010.506003.
Duckworth, A. L., Kirby, T. A., Gollwitzer, A., & Oettingen, G. (2013). From fantasy to action: Mental contrasting with implementation intentions (MCII) improves academic performance in children. *Social Psychological and Personality Science, 4*(6), 745–753. http://dx.doi.org/10.1177/1948550613476307.
Effeney, G., Carroll, A., & Bahr, N. (2013). Self-regulated learning: Key strategies and their sources in a sample of adolescent males1. *Australian Journal of Educational and Developmental Psychology, 13*, 58–74.
Evans, B. J., Baker, R. B., & Dee, T. S. (2016). Persistence patterns in Massive Open Online Courses (MOOCs). *Journal of Higher Education, 87*(2), 206–242. http://dx.doi.org/10.1353/jhe.2016.0006.
Friedman, J., Hastie, T., & Tibshirani, R. (2001). *The elements of statistical learning*. Berlin, Germany: Springer.
Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends, 59*(1), 64–71.
Guo, P. J., & Reinecke, K. (2014). Demographic differences in how students navigate through MOOCs. *Proceedings of the First ACM Conference on Learning @ Scale Conference - L@S '14*, 21–30. http://dx.doi.org/10.1145/2556325.2566247.
Hadwin, A. F., Nesbit, J. C., Jamieson-Noel, D., Code, J., & Winne, P. H. (2007). Examining trace data to explore self-regulated learning. *Metacognition and Learning, 2*(2–3), 107–124. http://dx.doi.org/10.1007/s11409-007-9016-7.
Hansen, J. D., & Reich, J. (2015). Democratizing education? Examining access and usage patterns in massive open online courses. *Science (New York, N.Y.),, 350*(6265), 1245–1248. http://dx.doi.org/10.1126/science.aab3782.
Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *The Behavioral and Brain Sciences, 33*(2–3), 61–83. http://dx.doi.org/10.1017/S0140525X0999152X. discussion 83–135.
Hew, K. F., & Cheung, W. S. (2014). Students' and instructors' use of massive open online courses (MOOCs): Motivations and challenges. *Educational Research Review, 12*, 45–58. http://dx.doi.org/10.1016/j.edurev.2014.05.001.
Ho, A. D., Chuang, I., Reich, J., Coleman, C. A., Whitehill, J., Northcutt, C. G., … Petersen, R. (2015). HarvardX and MITx : Two years of open online courses fall 2012-Summer 2014. *SSRN Electronic Journal, 10*, 1–37. http://dx.doi.org/10.2139/ssrn.2586847.
Hood, N., Littlejohn, A., & Milligan, C. (2015). Context counts: How learners' contexts influence learning in a MOOC. *Computers & Education, 91*, 83–91. http://dx.doi.org/10.1016/j.compedu.2015.10.019.
Jordan, K. (2014). Initial trends in enrolment and completion of massive open online courses. *The International Review Of Research In Open And Distributed Learning, 15*(1). http://dx.doi.org/10.19173/irrodl.v15i1.1651.
Karpicke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science, 319*(5865), 966–968. http://dx.doi.org/10.1126/science.1152408 (New York, N.Y.).
Kim, C., & Hodges, C. B. (2012). Effects of an emotion control treatment on academic emotions, motivation and achievement in an online mathematics course. *Instructional Science, 40*(1), 173–192. http://dx.doi.org/10.1007/s11251-011-9165-6.
Kizilcec, R. F. & Brooks, C. (in press). Diverse Big Data and Randomized Field Experiments in Massive Open Online Courses: Opportunities for Advancing Learning Research. In G. Siemens & C. Lang (Eds.), *Handbook on Learning Analytics & Educational Data Mining*.
Kizilcec, R. F., & Halawa, S. (2015). Attrition and Achievement Gaps in Online Learning. In *Proceedings of the Second ACM Conference on Learning @ Scale*. http://dx.doi.org/10.1145/2724660.2724680.
Kizilcec, R. F., Piech, C., & Schneider, E. (2013). Deconstructing Disengagement: Analyzing Learner Subpopulations in Massive Open Online Courses. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge*. http://dx.doi.org/10.1145/2460296.2460330.
Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. (2016). Recommending self-regulated learning strategies does not improve performance in a MOOC. In *Proceedings of the Third ACM Conference on Learning @ Scale*. http://dx.doi.org/10.1145/2876034.2893378.
Kizilcec, R. F., & Schneider, E. (2015). Motivation as a lens to understand online Learners: Toward data-driven design with the OLEI scale. *Transactions on Computer-Human Interactions (TOCHI), 22*(2). http://dx.doi.org/10.1145/2699735.
Koedinger, K. R., Mclaughlin, E. a, Kim, J., Zhuxin Jia, J., & Bier, N. L. (2015). Learning is Not a Spectator Sport: Doing is Better than Watching for Learning from a MOOC. In Proceedings of the Second ACM Conference on Learning @ Scale. http://dx.doi.org/10.1145/2724660.2724681.
Lajoie, S., & Azevedo, R. (2006). Teaching and learning in technology-rich environments. In *Handbook of educational psychology* (Vol. 2, pp. 803–821).
Lee, T.-H., Shen, P.-D., & Tsai, C.-W. (2008). Applying web-enabled problem-based learning and self-regulated learning to add value to computing education in Taiwan's vocational schools. *Educational Technology & Society, 11*(3), 13–25.
Liang, J.-C., & Tsai, C.-C. (2008). Internet self-efficacy and preferences toward constructivist internet-based learning environments: A study of pre-school teachers in Taiwan. *Educational Technology & Society, 11*(1), 226–237.
Liou, P.-Y., & Kuo, P.-J. (2014). Validation of an instrument to measure students' motivation and self-regulation towards technology learning. *Research in Science & Technological Education, 32*(2), 79–96. http://dx.doi.org/10.1080/02635143.2014.893235.
Little, R. J. A. (1988). Missing-data adjustments in large surveys. *Journal of Business & Economic Statistics, 6*(3), 287–296. http://dx.doi.org/10.2307/1391881.

Littlejohn, A., Hood, N., Milligan, C., & Mustain, P. (2016). Learning in MOOCs: Motivations and self-regulated learning in MOOCs. *The Internet and Higher Education, 29*, 40–48. http://dx.doi.org/10.1016/j.iheduc.2015.12.003.

Littlejohn, A., & Milligan, C. (2015). Designing MOOCs for profesional learners: Tools and patterns to encourage self_regulated learning. *Design Paper, 42*(June), 1–10. Retrieved from http://www.openeducationeuropa.eu/en/article/Design-Patterns-for-Open-Online-Teaching-and-Learning-Design-Paper-42-4.

Mullaney, T., & Reich, J. (2015). Staggered Versus All-At-Once Content Release in Massive Open Online Courses: Evaluating a Natural Experiment. *Proceedings of the Second (2015) ACM Conference on Learning @ Scale*, 185–194. http://dx.doi.org/10.1145/2724660.2724663.

Nawrot, I., & Doucet, A. (2014). Building engagement for MOOC students: introducing support for time management on online learning platforms. In Proceedings of the companion publication of the 23rd international conference on World wide web companion (pp. 1077–1082).

Nicol, D. J., & Macfarlane-dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education (2006), 31*(2), 199–218. http://dx.doi.org/10.1080/03075070600572090.

Niemi, H., Nevgi, A., & Virtanen, P. (2003). Towards self-regulation in web-based learning. *Journal of Educational Media, 28*(1), 49–71. http://dx.doi.org/10.1080/1358165032000156437.

Ogan, A., Walker, E., Baker, R., Rodrigo, M. M. T., Soriano, J. C., & Castro, M. J. (2015). Towards understanding how to assess help-seeking behavior across cultures. *International Journal of Artificial Intelligence in Education, 25*(2), 229–248. http://dx.doi.org/10.1007/s40593-014-0034-8.

Perna, L. W., Ruby, A., Boruch, R. F., Wang, N., Scull, J., & Evans, C. (2014). Moving through MOOCs: Understanding the progression of users in Massive Open Online Courses. *Educational Researcher*, (December), 421–432. http://dx.doi.org/10.3102/0013189X14562423.

Pintrich, P. R. (1999). The role of motivation in promoting and sustaining self-regulated learning. *International Journal of Educational Research, 31*(6), 459–470.

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451–502). San Diego: Academic Press. http://dx.doi.org/10.1016/B978-012109890-2/50043-3.

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review, 16*(4), 385–407.

Pintrich, P. R., & others. (1991). *A manual for the use of the motivated strategies for learning questionnaire (MSLQ)*.

Puustinen, M., & Pulkkinen, L. (2001). Models of self-regulated learning: A review. *Scandinavian Journal of Educational Research, 45*(3), 269–286.

Reich, J. (2014). MOOC completion and retention in the context of student intent. *EDUCASE Review Online*. Retrieved from http://er.educause.edu/articles/2014/12/mooc-completion-and-retention-in-the-context-of-student-intent.

Reich, J. (2015). Rebooting MOOC research. *Science, 347*(6217), 34–35.

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin, 138*(2), 353.

Rigotti, T., Schyns, B., & Mohr, G. (2008). A short version of the occupational self-efficacy scale: Structural and construct validity across five countries. *Journal of Career Assessment, 16*(2), 238–255.

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

Samruayruen, B., Enriquez, J., Natakuatoong, O., & Samruayruen, K. (2013). Self-regulated learning: A key of a successful learner in online learning environments in Thailand. *Journal of Educational Computing Research, 48*(1), 45–69.

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology, 19*(4), 460–475.

Schunk, D. H. (2005). Self-regulated learning: The educational legacy of Paul R. Pintrich. *Educational Psychologist, 40*(2), 85–94.

Seaton, D. T., Bergner, Y., Chuang, I., Mitros, P., & Pritchard, D. E. (2014). Who does what in a massive open online course? *Communications of the ACM, 57*(4), 58–65. http://dx.doi.org/10.1145/2500876.

Shah, D. (2015, December). MOOCs in 2015: Breaking down the numbers. *EdSurge*. Retrieved from https://www.edsurge.com/news/2015-12-28-moocs-in-2015-breaking-down-the-numbers.

Siadaty, M., Gašević, D., & Hatala, M. (2016). Associations between technological scaffolding and micro-level processes of self-regulated learning: A workplace study. *Computers in Human Behavior, 55*, 1007–1019. http://dx.doi.org/10.1016/j.chb.2015.10.035.

Sun, J. C.-Y., & Rueda, R. (2012). Situational interest, computer self-efficacy and self-regulation: Their impact on student engagement in distance education. *British Journal of Educational Technology, 43*(2), 191–204.

Taub, M., Azevedo, R., Bouchet, F., & Khosravifar, B. (2014). Can the use of cognitive and metacognitive self-regulated learning strategies be predicted by learners' levels of prior knowledge in hypermedia-learning environments? *Computers in Human Behavior, 39*, 356–367. http://dx.doi.org/10.1016/j.chb.2014.07.018.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society, Series B (Statistical Methodology), 58*(1), 267–288.

Tsai, C.-C. (2009). Conceptions of learning versus conceptions of web-based learning: The differences revealed by college students. *Computers & Education, 53*(4), 1092–1103.

Tsai, C.-C., Chuang, S.-C., Liang, J.-C., & Tsai, M.-J. (2011). Self-efficacy in internet-based learning environments: A literature review. *Educational Technology & Society, 14*(4), 222–240.

Wang, Y., & Baker, R. (2015). Content or platform: Why do students complete MOOCs? *MERLOT Journal of Online Learning and Teaching, 11*(1), 191–218.

Wang, C.-H., Shannon, D. M., & Ross, M. E. (2013). Students' characteristics, self-regulated learning, technology self-efficacy, and course outcomes in online learning. *Distance Education, 34*(3), 302–323.

Warr, P., & Downing, J. (2000). Learning strategies, learning anxiety and knowledge acquisition. *British Journal of Psychology, 91*(3), 311–333.

Weinstein, C. E., Acee, T. W., & Jung, J. (2011). Self-regulation and learning strategies. *New Directions for Teaching and Learning, 2011*(126), 45–53. http://dx.doi.org/10.1002/tl.443.

Winne, P. H. (2014). Issues in researching self-regulated learning as patterns of events. *Metacognition and Learning, 9*(2), 229–237. http://dx.doi.org/10.1007/s11409-014-9113-3.

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

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. *Metacognition in Educational Theory and Practice, 93*, 27–30.

Winne, P. H., & Nesbit, J. C. (2010). The psychology of academic achievement. *Annual Review of Psychology, 61*, 653–678.

Yukselturk, E., & Top, E. (2013). Exploring the link among entry characteristics, participation behaviors and course outcomes of online learners: An examination of learner profile using cluster analysis. *British Journal of Educational Technology, 44*(5), 716–728. http://dx.doi.org/10.1111/j.1467-8535.2012.01339.x.

Zheng, S., Rosson, M. B., Shih, P. C., & Carroll, J. M. (2015). Understanding Student Motivation, Behaviors and Perceptions in MOOCs. *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing - CSCW '15*, (pp. 1882–1895). http://dx.doi.org/10.1145/2675133.2675217.

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal, 45*(1), 166–183.

Zimmerman, B. J. (2015). *Self-regulated Learning: Theories, measures, and outcomes. International encyclopedia of the social & behavioral sciences*. Elsevier. Retrieved from http://www.sciencedirect.com/science/article/pii/B9780080970868260601.

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.) , *Handbook of self-regulation* (pp. 13–39). San Diego, USA: Academic Press. http://dx.doi.org/10.1016/B978-012109890-2/50031-7.

Zimmerman, B. J., & Pons, M. M. (1986). Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal, 23*(4), 614–628. http://dx.doi.org/10.2307/1163093.

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society, Series B (Statistical Methodology), 67*(2), 301–320. http://dx.doi.org/10.1038/203024b0.