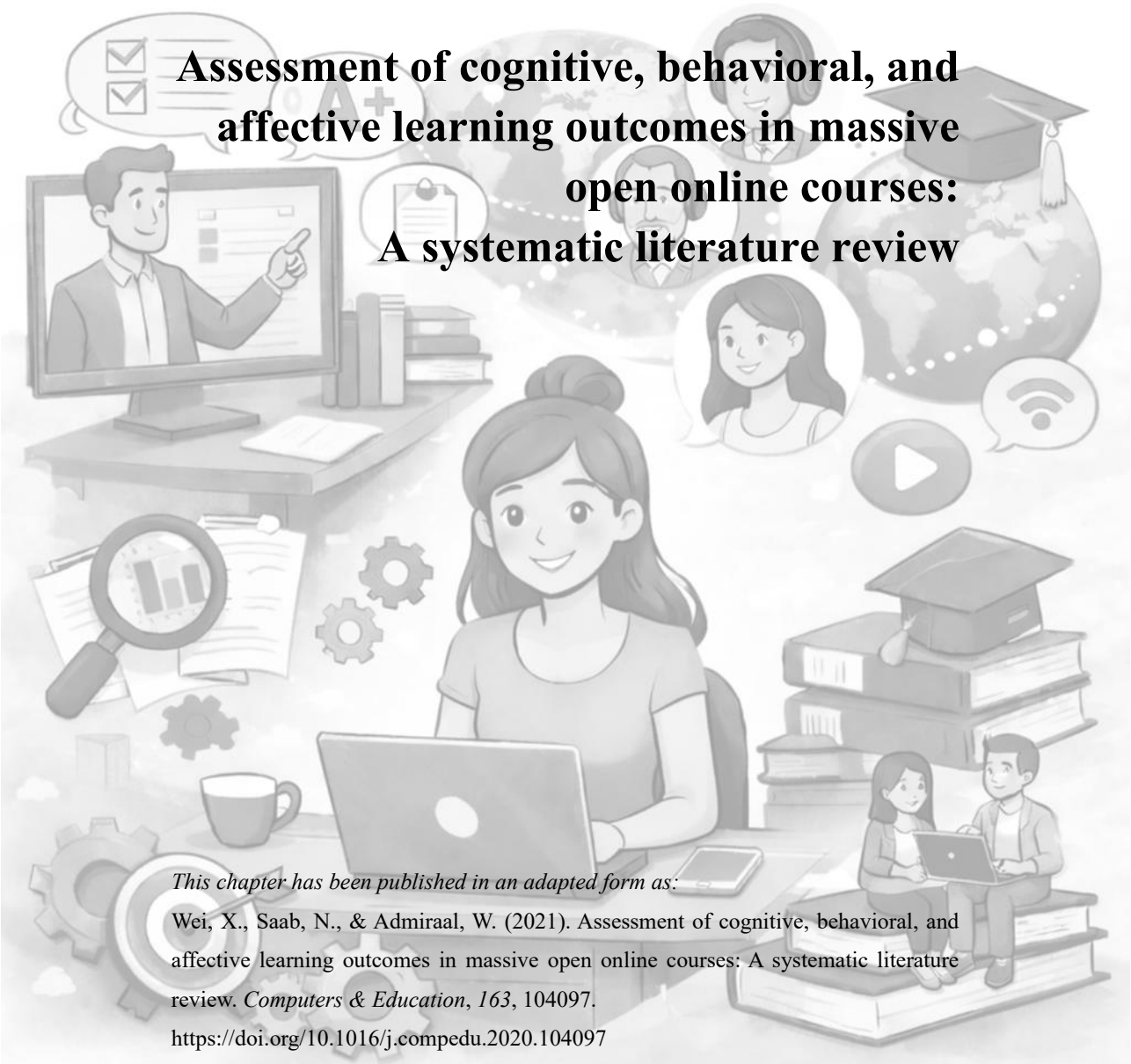
Assessment and learning engagement in massive open online courses

Wei, X.

Chapter 2

Assessment of cognitive, behavioral, and affective learning outcomes in massive open online courses: A systematic literature review



This chapter has been published in an adapted form as:

Wei, X., Saab, N., & Admiraal, W. (2021). Assessment of cognitive, behavioral, and affective learning outcomes in massive open online courses: A systematic literature review. *Computers & Education*, *163*, 104097. https://doi.org/10.1016/j.compedu.2020.104097

Abstract

This systematic review on massive open online courses (MOOCs) in higher education examined the research on the assessment of learning outcomes. This study aims to investigate the learning outcomes, related instruments, and assessment characteristics of these instruments in MOOCs. Learning outcomes that were examined in the studies reviewed concerned cognitive, behavioral, and affective learning outcomes, based on 65 peer-reviewed articles published between 2017 and 2019. There were 25 types of assessment approaches employed to examine these outcomes, and the assessment characteristics were identified. The results indicated that considering the assessment of learning outcomes at the beginning of course design could support the formulation of explicit assessment goals and in this way instructing learners to work towards learning outcomes. A combination of knowledge tests and skills tasks can be used to examine cognitive outcomes in a particular MOOC. Outcome-oriented feedback rubrics are beneficial to support learners' essay performance, and interpretations of the utilization of rubrics could better guide providers to give peer feedback. A variety of behavioral and affective outcomes reflect multiple aspects of participants learning in MOOCs, which might contribute to teachers better understand and support them to learn. Furthermore, assessment tasks throughout the course may differ in difficulty and complexity, which could align with different levels of motivation of the learners. The findings provide a holistic picture of learning outcomes and related assessment instruments in current MOOCs. Curriculum designers and teachers could benefit from this study to consider appropriate learning outcome variables and instruments to apply in their MOOC practices. Future research might investigate the motivation of learners to participate in a MOOC and how this changes during a MOOC. This could help MOOC designers and teachers to align how learners are motivated, what they want to learn, and what they actually do.

Keywords

Massive open online courses; learning outcomes; assessments; systematic review

2.1. Introduction

Massive open online courses (MOOCs) are a form of online learning with a large student population, and in most cases of free course offering and open access. The introduction of MOOCs has changed the way of delivering higher education resources. Although MOOCs originally were oriented towards informal learning, the acceptance of MOOCs as part of the formal campus education is more recently addressed (Hendriks et al., 2020). MOOCs, as campus-like course, have been realized in various practices ranging from undergraduate to graduate education (Yuan et al., 2014). Learners participate in MOOCs as part of blended learning or fully online learning, as an obligatory or an optional course, as a non-credit or a credit course, and as a complementary to or as a substitute for campus-based courses (Dale & Singer, 2019; Dandache et al., 2017; Littenberg-Tobias & Reich, 2020; Rayyan et al., 2016; Swinnerton et al., 2017). Learning assessment is one of the essential aspects of MOOCs. A variety of assessment methods naturally emerge to respond to the demand for measuring learning outcomes, in which computer grading tests and peer assessment are the common methods adopted to recognize and accredit students' learning (Admiraal et al., 2015).

Learning assessment in MOOCs has been regarded as a challenge, and both learners and teachers are aware that assessment instruments should be related to the learning goals of a particular MOOC (Yousef et al., 2014). Various learning outcomes in MOOCs have been studied, such as knowledge, skills, academic achievement, engagement, motivation, satisfaction, and perceptions of learning (Barak et al., 2016; Deng et al., 2020b; Jung et al., 2019; Jung & Lee, 2020; Li & Baker, 2018; Shapiro et al., 2017). Xiong and Suen (2018) have discussed different formats of assessments, such as discussion forums, Q&A sessions, quizzes, and automated grading essays. These assessment forms can be used as formative or summative assessments. In addition to automated tests, self-assessment and peer assessment (Admiraal et al., 2015; Huisman et al., 2018; Papathoma et al., 2015), other methods, such as groups of experts, semantic web, portfolio, and learning analytics, are recommended to be used as alternative assessments in MOOCs (del Mar Sánchez-Vera & Prendes-Espinosa, 2015). Although there is a debate about various methods that can be used as learning assessments in MOOC, in a wide range of disciplines in higher education, it lacks a focus on the connection between learning outcomes and assessment instruments.

As Sandeen (2013) claims, the assessment is the core consideration at the very beginning of the MOOC course design rather than a later add-on. The match between

learning goal, learning outcomes and assessment instruments is important to guarantee learners to realize the course required achievement (Bralić & Divjak, 2018; Ewais et al., 2020). Assessment of learning outcomes is key to measuring the actual achievements of the student, which is also critical for examining the effectiveness of instructional practices and students' learning (Alexandron et al., 2020; Anne Clark, 2002). In order to have holistic picture of learning outcomes and assessment instruments in higher education MOOCs in different disciplinary fields, the current review aims to obtain an overview of learning outcomes and related instruments in MOOCs.

2.2. Learning assessment in MOOCs

Learning assessment in MOOCs has been examined in previous review studies. Ebben and Murphy (2014) reviewed the nascent MOOC scholarship based on 25 peer-reviewed articles published from 2009 to 2013. They distinguished two main categories of MOOCs based on their pedagogical underpinning. The Connectivist MOOCs (cMOOCs), which are driven by the learning theory of Connectivism: openness, diversity, autonomy, and connected people within a network of diverse technologies can share and create knowledge (Bell, 2011; Downes, 2008). The pedagogical design of xMOOCs was similar to on-campus teacher-centered lectures delivering on the platforms. This period witnessed a change in research themes in MOOCs from studying engagement and creativity in early cMOOCs (2009-2012) to investigating the learning analytics and assessment in xMOOCs in later years (2012-2013). Learning assessment of MOOCs paid more attention to accreditation and grant to the credit, but focused less on assessment methodology itself. Several colleges and universities, like the University of Maryland and the University of Helsinki, have begun to grant the completion of a MOOC for credits. Moreover, automated assessment software, like Automated Essay Scoring, were applied in scoring writing assignments, but it was not valid to assess complicated essays.

Sa'don et al. (2014) employed data mining to extract the nascent research trend on MOOCs based on 164 papers published from January 2008 to July 2014. This review revealed that assessment and accreditation were included in a list of most popular research topics. Likewise, the assessment was also a focus of MOOC research addressed in the review on 146 empirical studies published between October 2014 and November 2016 (Zhu et al., 2018). This review implicated that researchers might more actively target the assessment practices of instructional designers and course instructors of MOOCs for future

research. Moreover, Pilli and Admiraal (2017) formulated suggestions for course design based on the review of 56 peer-reviewed studies related to learning outcomes in MOOCs published in or before 2015. These authors adopted the Biggs' 3P model to examine factors (e.g., students' characteristics, course features, and learning activities) influencing students learning outcomes (e.g., engagement, achievement) based on the extraction of previous scientific literature. They suggested to combine rubrics or other forms of guidance into peer and self-assessment, and to consider students' individual characteristic in assessment design. That same year, Vera et al. (2017) reviewed 38 peer-reviewed studies published between 2012 and 2016 and listed assessment types applied in MOOCs. Their findings revealed that tests and peer evaluation were common ways of assessment. In most cases, these two assessment approaches combined to assess learners' outcomes. Additionally, the review by Deng et al. (2019) analyzed 102 MOOC studies published from 2014 to 2016 that were concerned with teaching and learning. The findings show that the complexity of assessment methods is limited and the learning outcomes lacks diversity.

The above reviews synthesized the MOOC studies from 2009 to 2016, presenting a picture of scholarly attention on learning assessment in MOOCs. However, existing reviews are yet to have an in-depth insight into the learning outcomes and instruments in MOOCs. These reviews fail to reveal a comprehensive image of a variety of learning outcomes and the instruments itself in the discipline fields covered in MOOCs. The present study takes account of the categories of learning outcomes and related instruments, and the assessment characteristics of these instruments, providing an exhaustive analysis of the assessment of learning outcomes in recent MOOC studies. Therefore, the scientific knowledge base of the assessment of learning outcomes in MOOCs is reviewed by posing three research questions as follows:

RQ1: What learning outcomes are assessed in MOOCs?

RQ2: What instruments are employed to measure these learning outcomes?

RQ3: What assessment characteristics do these instruments have?

2.3. Methods

The systematic literature review is the methodology utilized in this study. The principles of both PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement and PRISMA explanation and elaboration were employed as the guidelines to conduct and report this review work (Liberati et al., 2009; Moher et al., 2009). The

eligibility criteria, information sources, search strategy, study records, data items, and data synthesis recommended in the methods part of the PRISMA-P checklist are described in the following sub-sections.

2.3.1. Search

The literature search was executed in all electronic databases available at the library of a research university in the Netherlands. This study aims to explore the assessment of learning outcomes in higher education MOOCs. Therefore, the search terms comprise the variants of the term MOOC and the synonyms of the term assessment, learning outcome, and higher education. Since the most identifiable features of MOOCs are massiveness and openness, we did not use other emerging types of MOOCs as the search terms, such as SPOCs, gMOOC, BOOCs, pMOOC, which are included in the taxonomy of MOOCs by Pilli and Admiraal (2016). The search criteria were used as follows. Search in the title: (MOOC OR massive open online course OR xMOOC or cMOOC). Search in any field: AND (assessment OR evaluation OR learning outcome OR performance OR achievement OR gain) AND (higher education OR college OR university). The last search was performed on 29 July 2019. Articles were included based on the following criteria: (a) peer-reviewed journal articles; (b) written in English; (c) include records with full access; (d) available in full text; and (e) published between 2017 and 2019. This review covered peer-reviewed articles published between 2017 to 2019 due to the previous work done by Chauhan (2014), Pilli and Admiraal (2017), and Vera et al. (2017). In total, the search yielded 248 articles, but only 230 articles were available in full-text from 12 electronic databases (see Table 2.1.).

2.3.2. Selection

There were a total of 230 articles available in full-text and downloaded for further manual selection. For each article, the title, abstract and full text were reviewed by adopting the exclusion criteria to determine the final samples included in the systematic review (see Table 2.2.). Figure 2.1. depicts the process of the literature selection, namely identification, screening, eligibility and analysis, and the number of articles changes after each of its stages. By checking the titles and abstracts and full texts, fifteen articles were removed due to the reasons of duplicates, written in languages other than English, and focus not on higher education, 215 articles remained for further selection. Eighteen articles, including reviews

and framework analysis, were filtered out because they did not provide the first-hand data relevant to the assessment of learning outcomes in MOOCs. After further reading on 197 articles, 132 articles were excluded since they did not report details related to learning outcomes and assessments in MOOCs. Ten percent of 230 articles were rated by a co-author and there was a 100% match in the inclusion and exclusion of articles by both raters. Ultimately, the manual selection resulted in 65 eligible articles for the systematic review.

Table 2.1. Databases collection.

Name of databases	Number of articles found in the databases
ScienceDirect (Elsevier)	50
ProQuest	49
Educational Resources Information Center (ERIC)	31
SpringerLink	28
Web of Science	20
Google Scholar	18
Taylor & Francis Online	16
SAGE Journals	8
Wiley Online Library	5
JSTOR (Journal STORage)	3
Scholarly and Academic Information Navigator (CiNii)	1
PubMed	1
Total	**230**

Table 2.2. Number of excluded articles based on the exclusion criteria.

Exclusion criteria	Number of articles excluded based on information in	
	Title/abstract	Full text
Review articles, framework analysis	16	2
Written in languages other than English	11	
No reported details related to learning outcomes and assessments in MOOCs	107	25
Focus not on higher education	1	
Duplicates	3	

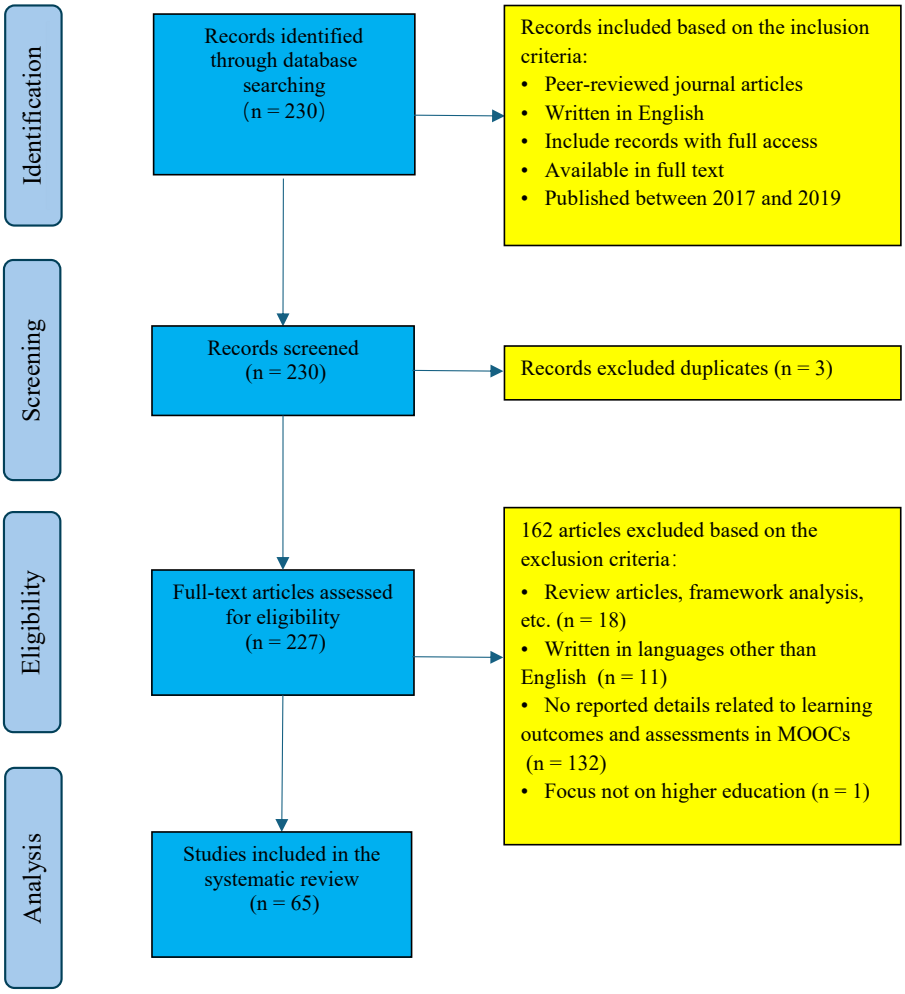


Figure 2.1. Flowchart of the study selection process following the guidelines of the PRISMA (Moher et al., 2009).

2.3.3. Data analysis

The details related to learning outcomes and assessment instruments are depicted in a matrix. Based on the findings of the selected articles, the learning outcomes were categorized by adopting the taxonomy of Bloom and Krathwohl as cognitive outcomes (e.g. knowledge and skills), behavioral outcomes (e.g. actual behavior), and affective outcomes (e.g. value, attitudes, and motivation; Bloom et al., 1956; Krathwohl et al., 1964). As subcategories per main outcome category we used the categories Post et al. (2019) have

employed in their review study to specify learning outcomes. The learning outcomes and instruments included in the studies reviewed were examined. In order to group the assessment methods into types of assessment that were used to measure the learning outcomes (i.e. quiz, assignment, survey, user data), we applied the analysis method of Day et al. (2018). Each article has been coded by (a) assessment type, (b) author/date, (c) country/region, (d) discipline, (e) purpose, (f) assessor, (g) frequency, (h) feedback, (i) mandatory (i.e., learners have to take the assessments), (j) scoring (i.e., contribution to the percentage of the final grades), and (k) reward (i.e., to obtain credits/certificate).

2.4. Results

In Table 2.3, we have included an overview of the articles selected with respect of the learning outcomes these address and the instruments these use. In Tables 2.4, 2.5 and 2.6, we have included the assessment characteristics each article address with respect to cognitive, behavioral and affective learning outcomes, respectively.

2.4.1. Learning outcomes assessed in MOOCs

2.4.1.1. Cognitive outcomes

Cognitive outcomes relate to the acquisition of knowledge and information and intellectual skills (Bloom, 1956).

Knowledge

In thirty-two studies, learners' conceptual knowledge and academic achievement were measured as the outcomes of MOOCs. *Conceptual knowledge* refers to knowledge and information that learners have understood and obtained from the subject matter (Bloom, 1956). For example, foundational knowledge in psychology, statistical methods, and computer programming (Najafi et al., 2017), knowledge concerned the new computing technologies and business processes integration (Qiu & Xu, 2018), and content knowledge relevant to the astronomy (Formanek et al., 2017) and dementia (Eccleston et al., 2019) were assessed. Learners' academic achievement in MOOCs that was measured in courses such as Biology (Kahan et al., 2017), Algebra and Pre-Calculus (Li & Baker, 2018), and Classical Sociological Theory (van de Oudeweetering & Agirdag, 2018).

Intellectual skills

Eight studies reported intellectual skills that learners acquired from MOOCs. *Intellectual skills* concerned a set of domain-specific and generic abilities involved in the thinking processes, such as reasoning, comprehension, thinking abilities, problem-solving, decision-making skills (Shavelson & Huang, 2003). In the reviewed studies, learners' professional capabilities and broad abilities were assessed. For example, in courses about the physiotherapy management of spinal cord injuries in Harvey et al. (2017), learner clinical reasoning skills were assessed. In the study by Najafi et al. (2017), computational thinking in a programming course and critical thinking in a psychological course were identified as learning outcomes of course participants. We see that Meek et al. (2017) examined learner writing proficiency of written summary in a MOOC of biomedical science. Participants in this course showed different levels of proficient outcomes (pass, grades A-C; fail, grades D and E) in written summary exercises. Similarly, in the study of Comer (2017), the MOOC titled English Composition measured learner writing-related skills by four writing projects. Moreover, broad abilities concerned self-regulated learning were measured in the cases of Onah and Sinclair (2017), referring to goal settings, task strategies, time management, environment structuring, help seeking, and self-evaluation.

2.4.1.2. Behavioral outcomes

Behavioral outcomes indicate to what extent learners engage in learning, based on their behavior in learning activities (e.g. watching video lectures, contributing to discussion forums, and completing assessments), which show positive association with course achievement and completion (Kahan et al., 2017). In terms of behavioral outcomes in the studies reviewed, engagement in learning activities and course completion have been examined to illustrate the observable behavior of learners in MOOC learning.

Engagement

Thirty-nine studies measured learners' course engagement (i.e., engagement in course content, course assessment, and discussion forums) during MOOC learning (Balint et al., 2017; Bonafini, 2017; Chang et al., 2019; Conijn et al., 2018; Duan et al., 2019; Jia et al., 2019; Jung & Lee, 2018; Khalil & Ebner, 2017; Lan et al., 2019; Luo et al., 2018; Patterson, 2018; Torres & Beier, 2018; Wong, Khalil, et al., 2019; Yang & Su, 2017). For example, Kahan et al. (2017) investigated the number of video lectures viewed/downloaded, video

questions answered, exams/quizzes submitted, threads views/opened and posts/comments written in the discussion forums, which were used to determine the participant's engagement. Based on these the authors identify seven types of participant behavior, namely tasters, downloaders, disengagers, offline engagers, online engagers, moderately social engagers, and social engagers. Likewise, Almeda, Zuech, Baker, Utz, Higgins, & Reynolds (2018), Bonafini et al. (2017), and Chiu and Hew (2018) focused on the number of posts/replies/comments, and the duration of interactions in the discussion forums. In Kizilcec et al. (2017), learner interaction with course materials and the time-cost per-session activity were assessed, such as lecture begin/complete/revisit, and assessment attempt/pass/revisit. Finally, Kovanović et al. (2019) explored seven aspects of students' course engagement: course access, completed assignments, video lectures watched, course navigation, discussion access, discussion contribution, and discussion reputation.

Course completion

The outcomes related to course completion were reported in sixteen studies, and earning the course certificate were addressed in fourteen studies (Khalil & Ebner, 2017; Lan et al., 2019; Lee, 2019; Loizzo et al., 2017; Singh & Mørch, 2018; Torres & Beier, 2018; L. Wang et al., 2018). Since the requirements of course completion and course certificate differ in each MOOC, the differentiated degrees of completion could lead learners to receive different types of certificates in that course. Bonafini et al. (2017) reported one course via the Coursera platform was offering two options for a certificate of completion. In this course, the submission of all six weekly tasks is mandatory to obtain a Statement of Accomplishment certificate. When learners would like to obtain a Statement of Accomplishment with Distinction certificate, they also had to submit 12 additional peer reviews. Some studies, such as Najafi et al. (2018), Jia et al. (2019), Sunar et al. (2018), and Patterson (2018), also reported the data on course completion and course certificates.

2.4.1.3. Affective outcomes

Affective outcomes refer learners' perceptions of their learning in MOOCs. To identify the aspects of affective learning gains in MOOCs, learners' course satisfaction (i.e., positive evaluation of courses), perceptions of learning experiences (i.e., learners' perceived appreciation of course learning), and perceptions of learning benefits (i.e., learners' perceived learning improvement) have been examined (Post et al., 2019).

Course satisfaction

The satisfaction revealed participants' positive evaluation of the MOOCs that was measured in seven studies (Khalil & Ebner, 2017; Li & Moore, 2018; Valdivia Vázquez et al., 2018). For example, De Lima & Zorrilla (2017) measured learners' satisfaction concerning the realization of the learning goals. The results show that learners expressed high satisfaction with peer interaction, such as supporting and commenting by peers, and resource sharing by other participants. Li (2019) and Rabin et al. (2019) assessed learners' overall satisfaction with the courses. Another study by Watson et al. (2017) investigated the satisfaction of MOOC learners with more specific aspects of the course. Course participants reported their satisfaction with course learning, learning activities, and assignments. This study carried out comparison among different learner groups with respect to their course satisfaction. The comparative analysis indicated that males reported significantly lower course satisfaction than females, and learners who tended to prefer instructor-led activities had significantly lower satisfaction than learners who tended to prefer exploratory learning activities.

Perceptions of MOOCs experiences

Nineteen studies investigated learners' experiences of learning in MOOCs. Several studies explored emotional feelings of course participants in MOOCs. For example, Buhr et al. (2019) reported learners who were bored with the course or enjoyed it, and Loizzo et al. (2017) investigated participants' enjoyment gained within MOOC. In the study of L. Wang et al. (2018), eight types of learners' emotional tendencies were presented, namely happy, sad, angry, disappointed, surprised, proud, in love, and scared. Some studies investigated participants' perceptions of course learning (e.g., Jung et al., 2019; Li & Moore, 2018; Meek et al., 2017). For example, both Poquet et al. (2018) and Kovanović et al. (2019) employed the Community of Inquiry (CoI) model to focus on issues related to the social presence (i.e., group cohesion, open communication, and affective expression) of learners in the MOOC environment. Poquet et al. (2018) evaluated the social presence of the different learner groups (i.e., regular posters, occasional posters, and non-posters) in the discussion forums. Kovanović et al. (2019) examined the social presence of three student clusters in terms of course engagement and discussion activity (i.e., limited users, selective users, broad users).

In some studies, learners reported their perceptions of course design, course content,

teaching presence, ease of use and usefulness of the course, course usability, network complementarity/size, ease or difficulty of behavioral control, and user preference (e.g., Jung et al., 2019; Jung & Lee, 2018; Kovanović et al., 2019; Li et al., 2018; Rabin et al., 2019; Singh & Mørch, 2018; Yang & Su, 2017). Moreover, course participants expressed challenges and barriers that they were confronted with in MOOCs. According to the studies of Shapiro et al. (2017) and Watson et al. (2017), learners were facing obstacles concerned with the quite substantial course workload, lack of deep conversation in discussion forums, inadequate prior knowledge of the topics, previous terrible experience with the subjects, and no enough time for learning. Meanwhile, some learners had pointed out that the course content was not explained explicitly and deep enough (Singh & Mørch, 2018; Swinnerton et al., 2017).

Perceptions of MOOCs benefits

Learning benefits that learners have reported were addressed in twenty-five studies. Some studies investigated course participants' motivation (e.g., Loizzo et al., 2017; Valdivia Vázquez et al., 2018) and self-efficacy to MOOCs (e.g., Jung & Lee, 2018). For example, Shapiro et al. (2017) found learners' diverse motivation to attend a MOOC, such as related to knowledge augmentation, improvement of work abilities, development of future career, and personal interests. DeBoer et al. (2019) conducted a randomized control trial to examine students' self-efficacy in a MOOC about neuroscience topics. At the end of this MOOC, the treatment group students have a significant growth in self-efficacy in the post-test compared with the results of the pre-test, and the growth of self-efficacy was larger than students in the control group. Both Watson et al. (2017) and Hudson et al. (2019) investigated learners' perceptions of attitudinal learning (i.e. affective, behavioral, and cognitive learning) in MOOCs. Furthermore, several studies measured other perceived benefits of MOOCs, such as general learning effectiveness (Jung et al., 2019; Swinnerton et al., 2017), knowledge sharing and increasing (Singh & Mørch, 2018; Stich & Reeves, 2017), study skills improvement (Rodriguez & Armellini, 2017), sense of progress (Jung et al., 2019), confidence (Bonafini, 2017; Li & Moore, 2018), and user persistence (Li et al., 2018; Najafi et al., 2018).

Table 2.3. An overview of learning outcomes and assessment instruments in the studies reviewed (n = 65).

Authors (year)	Cognitive outcomes		Behavioral outcomes		Affective outcomes	
	Outcome type	Instruments	Outcome type	Instruments	Outcome type	Instruments
1. Almeda et al. (2018)	K	Course requirements	E	User data		
2. Balint et al. (2017)			E	User data		
3. Bonafini (2017)	K; IS	Discussion forum	E	User data	PB	Self-assessment
4. Bonafini et al. (2017)			E; COM	User data		
5. Buhr et al. (2019)					PE	Survey
6. Chang et al. (2019)	K	Group embedded figure test	E	Group embedded figure test		
7. Chiu & Hew (2018)	K	Quiz, assignment	E	User data		
8. Comer (2017)	K; IS	Writing projects			PB	User data
9. Conijn et al. (2018)	K	Quiz, exam, assignment	E	User data		
10. de Lima & Zorrilla (2017)			E	User data	SAT	Survey
11. DeBoer et al. (2019)	K	Exam			PB	Survey
12. Duan et al. (2019)			E	User data		
13. Eccleston et al. (2019)	K	Survey				
14. Emanuel & Lamb (2017)	K	Assessment	E; CER	User data		
15. Formanek et al. (2017)	K	Quiz, assignment, online activities				

Table 2.3. An overview of learning outcomes and assessment instruments in the studies reviewed (n = 65).

Authors (year)	Cognitive outcomes		Behavioral outcomes		Affective outcomes	
	Outcome type	Instruments	Outcome type	Instruments	Outcome type	Instruments
16. Harvey et al. (2017)	K; IS	Assessment	E	User data		
17. Hudson et al. (2019)			E	Survey	PB	Survey, interview
18. Jia et al. (2019)	K	Test, exam, discussion forum	E; COM	User data		
19. Jung & Lee (2018)			E	Survey	PB; PE	Survey
20. Jung et al. (2019)					PB; PE	Survey
21. Kahan et al. (2017)	K	Quiz, exam	E; CER	User data		
22. Khalil & Ebner (2017)	K	Quiz, practical work	E; COM; CER	User data	SAT	Survey
23. Kizilcec et al. (2017)			E; CER	Survey, user data		
24. Kormos & Nijakowska (2017)					PB; PE	Survey
25. Kovanović et al. (2019)	K; IS	Homework, labs	E	User data	PB; PE	Survey
26. Krasny et al. (2018)					PB; PE	Survey, interview
27. Lan et al. (2019)	K	Assessment	E; COM; CER	User data		
28. Lee (2018)	K	Checkpoints, quiz, homework, exam	E; CER	User data		
29. Lee (2019)	K	Checkpoints, homework problems, quiz, exam	E; COM; CER	User data		

Table 2.3. An overview of learning outcomes and assessment instruments in the studies reviewed (n = 65).

Authors (year)	Cognitive outcomes		Behavioral outcomes		Affective outcomes	
	Outcome type	Instruments	Outcome type	Instruments	Outcome type	Instruments
30. Li & Baker (2018)	K	Quiz	E	User data		
31. Li & Moore (2018)	K	Quiz, exam, problem sets			PB; PE; SAT	Survey, interview
32. Li (2019)					PB; PE; SAT	Survey
33. Li et al. (2018)			COM	Survey	PB; PE	Survey
34. Loizzo et al. (2017)			COM; CER	Interview	PB; PE	Interview, observation
35. Luo et al. (2018)			E	User data		
36. Maldonado-Mahauad et al. (2018)			E	User data		
37. Meek et al. (2017)	K; IS	Assignment	E	Survey	PB; PE	Survey
38. Moreno-Marcos et al. (2018)	K	Test, assignment				
39. Najafi et al. (2017)	K; IS	Essay, quiz, assignment, exercise, exam, test, discussion forum				
40. Najafi et al. (2018)	K	Quiz, assignment, course project	COM	User data	PB	Survey

Table 2.3. An overview of learning outcomes and assessment instruments in the studies reviewed (n = 65).

Authors (year)	Cognitive outcomes		Behavioral outcomes		Affective outcomes	
	Outcome type	Instruments	Outcome type	Instruments	Outcome type	Instruments
41. Onah & Sinclair (2017)	IS	Survey				
42. Patterson (2018)	K	Exam, quiz, assignment	E; COM; CER	User data		
43. Poquet et al. (2018)			E	User data	PE	Survey
44. Qiu & Xu (2018)	K	Assignment, exam	CER	Survey		
45. Rabin et al. (2019)			E; CER	User data	PB; PE; SAT	Survey
46. Rizvi et al. (2019)			E	User data		
47. Rodriguez & Armellini (2017)					PB	Survey
48. Shapiro et al. (2017)					PB; PE	Interview
49. Singh & Morch (2018)	K	Quiz, assessment	E; COM	Survey, observation	PB; PE	Survey
50. Stich & Reeves (2017)			COM	Survey	PB; PE	Survey
51. Sun & Bin (2018)	IS	User data				
52. Sunar et al. (2018)			E; COM	User data		
53. Swinnerton et al. (2017)			E	Survey	PB; PE	Survey
54. Torres & Beier (2018)	K	Quiz, discussion forum	E; COM	User data		

Table 2.3. An overview of learning outcomes and assessment instruments in the studies reviewed (n = 65).

Authors (year)	Cognitive outcomes		Behavioral outcomes		Affective outcomes	
	Outcome type	Instruments	Outcome type	Instruments	Outcome type	Instruments
55. van de Oudeweetering & Agirdag (2018)	K	Assignment, quiz, exam	COM	Survey		
56. van den Beemt et al. (2018)	K	Quiz, assignment	E; CER	User data		
57. Vázquez et al. (2018)					PB; SAT	Survey
58. Wang, Anderson & Chen (2018)			E	User data		
59. Wang, Hu & Zhou (2018)			E; COM; CER	User data	PE	User data
60. Watson et al. (2017)	K	Assessment, quiz, exam, exercise	CER	User data	PB; SAT	Survey, interview
61. Williams et al. (2018)			E	User data		
62. Wong et al. (2019)			E	User data		
63. Wu et al. (2019)					PB	Interview
64. Xing (2019)	K	Quiz, test	COM	User data		
65. Yang & Su (2017)			E	Survey	PB; PE	Survey

Note: K = knowledge, IS = intellectual skills, E = engagement, COM = course completion, CER = course certificate, PE = perceptions of MOOCs experience, PB = perceptions of MOOCs benefits, SAT = course satisfaction.

2.4.2 Instrument types and characteristics

2.4.2.1. Assessment instruments of cognitive outcomes

As displayed in Table 2.4, a total of twenty types of assessment instruments were adopted to measure cognitive outcomes. All these measurements refer to a knowledge test, in which quizzes, assignments, tests, exams, and exercises were the most common (e.g., Chiu & Hew, 2018; van den Beemt et al., 2018; Watson et al., 2017). Only eleven types were employed to assess intellectual skills of learners acquired in MOOCs, referring to assignments, surveys, assessments, discussion forums, tests, exercises, user data, essays, labs, and writing projects (e.g., Comer, 2017; Krasny et al., 2018). In the discipline column of Table 2.4, we see that quizzes, exams, and assignments were widely used types across different disciplines to examine knowledge outcomes of learners mastered in MOOCs. Concerning some specific disciplines, such as Computer Science, Social Science, and Humanities, the instruments used focused more on skills and competencies. For example, critical thinking and argumentation were measured with essays in psychological courses, and computational thinking was assessed with programming exercises in programming courses (Najafi et al., 2017).

In Table 2.4, we found that the assessment approaches can be adopted for formative or summative assessment purposes. Several studies revealed that summative measurements, like mid-term exams and final exams, aimed to examine learners' academic achievement (DeBoer et al., 2019; Lee, 2019; Patterson, 2018). Najafi et al. (2017) reported formative assessment approaches with non-graded in-video quizzes and exercises contributed to learners mastering course content in a MOOC about computer programming.

In the assessor column of Table 2.4, cognitive outcomes were assessed by the computer, learner-self, peer, instructor, researcher, or it was not reported who performed the assessment. In some cases, the computer automatically graded the submissions (e.g. tests, quizzes, assignments, exams, checkpoints, and exercises), and learners could receive automatic grades (Khalil & Ebner, 2017; Lee, 2018). Researchers employed other measurements, such as learners' self-reports, answered questions in interviews, and learning tracked by user data, to investigate how well learners performed in the course learning. For example, in some studies learners reported that they gained some benefits in MOOCs, such as the development of self-regulated learning skills, pedagogical skills, organizational skills, leadership in the community, programming

ability, and comprehension ability (Krasny et al., 2018; Onah & Sinclair, 2017; Sun & Bin, 2018; Watson et al., 2017). In the cases of peer review, peers and instructors graded assignments and essays by rubrics and gave feedback that guided learners to make revisions to coursework (Comer, 2017; Formanek et al., 2017).

The amount of assessments learners were facing was quite different across studies. The frequency column in Table 2.4 shows these discrepancies. The number of assessments is flexible and generally depends on the exact duration of courses. The majority of studies describe the number of assessments embedded in one MOOC ranges from one to ten, like quizzes, exams, tests, exercises, labs, and assignments. However, Patterson (2018) reported the highest number of assessments in one MOOC: learners had to participate in 53 short quizzes. In contrast, learners in a neuroscience MOOC only had to participate in one graded final exam at the end of the semester (DeBoer et al., 2019). Additionally, the assessments also occurred weekly or bi-weekly, such as weekly quizzes and optional fortnightly advanced problem set in a MOOC about chemistry (Li & Moore, 2018).

Concerning feedback, nine studies reported instruments like essays, writing assignments, and writing projects, which provide some types of feedback to course participants. Rubrics are used in some measurements to guide assessors to give feedback on learners' coursework. For example, research by Formanek et al. (2017) focused on how peer assessment contributed to learners' learning and motivation in a MOOC about astronomy. In this study, peers gave feedback by employing a global 0-4 points scale rubrics in the first peer-reviewed writing assignment, which aims to evaluate whether learners correctly and concisely answered the questions. In the second peer-reviewed writing assignment, a more exhaustive rubric was applied as guideline for learners' self-assessment and peer-assessment. Comer (2017) conducted a qualitative study to investigate learners' learning outcomes grounded on peer feedback in an English composition course. The elaborate rubrics demonstrated learning outcomes items in writing projects. These rubrics guided learners to provide feedback in two stages: (1) drafting and formative feedback, and (2) revision and peer evaluation. Learners were encouraged to provide specific feedback that could support peers to work toward the courses learning outcomes. In addition, web-based tests, a kind of complicated automatic-grading form, provided alternative answers as feedback to learners (Najafi et al., 2017).

As to the columns of mandatory, scoring, and reward in Table 2.4, in most

MOOCs learners were required to participate in course assessments to gain a final course grade and to obtain credits and certificates (e.g., Comer, 2017; Kahan et al., 2017; Lee, 2019). The majority of studies described the proportion of graded assessments of final grade ranging from 8% to 100% (e.g. quizzes, assignments, exams, discussion forums, homework, essays, and writing projects). Once learners reached the passing requirements, they would receive the certificate or credits. As described in the studies reviewed, the courses differ in the reward standards. For example, learners had to receive a score of at least 60% of the final grade to obtain a completion certificate, and a score of at least 90% of the final grade to acquire a distinction certificate (Patterson, 2018). In a MOOC on biology, learners must score at least 78% of the course grade to receive course credits (Kahan et al., 2017). Yet, in a course about computer science, learners had to score at least 75% of grade in every single trail and to complete additional practical work, in order to receive credits (Khalil & Ebner, 2017).

Table 2.4. An overview of assessment characteristics of cognitive outcomes in the studies reviewed.

Assessment type	Authors (year)	Country/ Region	Discipline	Cognitive Outcomes	Assessment characteristics						
					Purpose	Assessor	Frequency	Feedback	Mandatory	Scoring	Reward
Quiz	7. Chiu & Hew (2018)	China-Hong Kong	No info	K	No info	No info	2	No info	Yes	No info	Certificate
	9. Conijn et al. (2018)	The Netherlands	Computer Science	K	No info	No info	9 graded quizzes; 5 practice quizzes	No info	No info	No info	No info
	15. Formanek et al. (2017)	USA	Physical Science & Engineering	K	No info	Computer	13	No info	Yes	60% of grade	No info
	21. Kahan et al. (2017)	Israel	Biology	K	Formative	No info	6	No info	Yes	50% of grade	Credit, certificate
	22. Khalil & Ebner (2017)	Austria	No info	K	Formative	Computer	10	No	Yes	No info	Credit, certificate
	28. Lee (2018)	USA	Engineering	K	Summative	No info	No info	No info	Yes	36% of grade	Certificate
	29. Lee (2019)	USA	Engineering	K	Summative	Computer	No info	No info	Yes	36% of grade	Certificate
	30. Li & Baker (2018)	USA	Mathematics	K	No info	No info	32 in Algebra and Pre-calculus 2, respectively; 52 in Pre-calculus	No info	Yes	No info	Certificate
	31. Li & Moore (2018)	USA	Chemistry	K	No info	No info	Weekly quizzes	No info	No info	No info	Certificate
	39. Najafi et al. (2017)	Canada	Health, Computer Science, Social Science & Humanities	K	Formative	Computer	307 and 140 in-video quizzes in 3 hard MOOCs), and 3 soft MOOCs, respectively; 2 quizzes in AboriginalEd; 1 bonus quiz in Mental Health	No info	Yes	Not contribute to grade; 40% of grade in AboriginalEd; 10% of grade in Mental Health.	No info
	40. Najafi et al. (2018)	Canada	Computer Science	K	No info	No info	≥2 in Bio1 course; 5 in iOSApp1 and iOSApp2, respectively.	No info	Yes	No info	No info
	42. Patterson (2018)	USA	No info	K	No info	No info	53	No info		10% of grade	Certificate

Table 2.4. An overview of assessment characteristics of cognitive outcomes in the studies reviewed.

Assessment type	Authors (year)	Country/ Region	Discipline	Cognitive Outcomes	Assessment characteristics						
					Purpose	Assessor	Frequency	Feedback	Mandatory	Scoring	Reward
	49. Singh & Morch (2018)	Norway	Nature & Environment	K	No info	No info	No info	No info	No info	No info	No info
	54. Torres & Beier (2018)	USA	No info	K	No info	Instructor	No info	No info	Yes	60% of grade	Certificate
	55. van de Oudeweetering & Agirdag (2018)	The Netherlands	Physical Science & Engineering, Social Science, Data Science	K	No info	No info	No info	No info	Yes	No info	No info
	56. van den Beemt et al. (2018)	The Netherlands	Data Science	K	No info	No info	6 weekly quizzes; 1 final quiz	No info	Yes	No info	Certificate
	60. Watson et al. (2017)	USA	Social Science	K	Formative	No info	10	No info	No info	No info	No info
	64. Xing (2019)	China-Taiwan	Business & Management, Computer Science, Education, Humanities, Mathematics & Statistics, the Medical Science, the Physical Science, the Professional & Applied science, interdisciplinary and other fields.	K	No info	No info	No info	No info	No info	No info	No info
Assignment	7. Chiu & Hew (2018)	China-Hong Kong	No info	K	No info	Peer	4	No info	Yes	No info	Certificate
	9. Conijn et al. (2018)	The Netherlands	Computer Science	K	No info	Peer	1 Peer-reviewed assignment	Yes	Yes	No info	No info
	15. Formanek et al. (2017)	USA	Physical Science & Engineering	K	Formative	Peer, instructor	3	Rubrics	Yes	20% of grade	No info

Table 2.4. An overview of assessment characteristics of cognitive outcomes in the studies reviewed.

Assessment type	Authors (year)	Country/ Region	Discipline	Cognitive Outcomes	Assessment characteristics						
					Purpose	Assessor	Frequency	Feedback	Mandatory	Scoring	Reward
	37. Meek et al. (2017)	UK	Health & Psychology	K	Summative	Peer, academics	1	Rubrics	Yes	No info	No info
	38. Moreno-Marcos et al. (2018)	Spain	Computer Science	K	Formative	Peer	2	Rubrics	Yes	25% of grade	No info
	39. Najafi et al. (2017)	Canada	Computer Science	IS	Formative	Peer, computer	3 assignments in Programming1; 2 assignments in Programming2 and Stats, respectively.	Rubrics	Yes	40% of grade in Programming1; 40% of grade in Programming2; 30% of grade in Stats.	No info
	40. Najafi et al. (2018)	Canada	Computer Science	K	No info	No info	1 in Bio1 course	No info	Yes	No info	No info
	42. Patterson (2018)	USA	No info	K	No info	No info	9	No info	Yes	45% of grade	Certificate
	44. Qiu & Xu (2018)	USA	Management & Leadership	K	No info	Computer	No info	No info	Yes	Yes	Certificate
	55. van de Oudeweetering & Agirdag (2018)	The Netherlands	Physical Science & Engineering, Social Science, Data Science	K	No info	No info	No info	No info	Yes	No info	No info
	56. van den Beemt et al. (2018)	The Netherlands	Data Science	K	No info	Peer	1	No info	Yes	No info	Certificate
Exam	9. Conijn et al. (2018)	The Netherlands	Computer Science	K	No info	No info	1 final exam	No info	No	No info	No info
	11. DeBoer et al. (2019)	USA	Neuroscience, Engineering	K	Summative	No info	1 final exam	No info	Yes	100% of grade	Certificate
	18. Jia et al. (2019)	China	Medical Health	K	No info	No info	1 final exam	No info	Yes	50% of grade	No info
	21. Kahan et al. (2017)	Israel	Biology	K	Summative	No info	1 final exam	No info	Yes	50% of grade	Credit, certificate
	28. Lee (2018)	USA	Engineering	K	Summative	No info	1 midterm exam; 1 final exam	No info	Yes	23% of grade	Certificate
	29. Lee (2019)	USA	Engineering	K	Summative	Computer	1 midterm exam; 1 final exam	No info	Yes	23% of grade	Certificate
	31. Li & Moore (2018)	USA	Chemistry	K	No info	No info	1 midterm exam; 1 final exam	No info	No	No info	Certificate
	39. Najafi et al. (2017)	Canada	Computer Science	K	Summative	Computer	1 final exam	No info	Yes	50% of grade	No info
	42. Patterson (2018)	USA	No info	K	Summative	No info	1 final exam	No info	Yes	45% of grade	Certificate

Table 2.4. An overview of assessment characteristics of cognitive outcomes in the studies reviewed.

Assessment type	Authors (year)	Country/ Region	Discipline	Cognitive Outcomes	Assessment characteristics						
					Purpose	Assessor	Frequency	Feedback	Mandatory	Scoring	Reward
	44. Qiu & Xu (2018)	USA	Management & Leadership	K	No info	Computer	1 final exam	No info	Yes	No info	Certificate
	55. van de Oudeweetering & Agirdag (2018)	The Netherlands	Physical Science & Engineering, Social Science, Data Science	K	No info	No info	No info	No info	Yes	No info	No info
	60. Watson et al. (2017)	USA	Social Science	K	Summative	No info	1 midterm exam; 1 final exam	No info	Yes	No info	Certificate
Survey	13. Eccleston et al. (2019)	Australia	No info	K	No info	Self	1 pre-survey; 1 post-survey	No	No	No	No
	41. Onah & Sinclair (2017)	UK	Computer Science	IS	No info	Self	1	No	No	No	No
Assessment	14. Emanuel & Lamb (2017)	USA	Art & Culture	K	Formative	Computer	No info	No info	Yes	50% of grade or greater	Certificate
	16. Harvey et al. (2017)	Australia	Healthcare	K; IS	No info	Self	Self-assessments; the pre-MOOC and post-MOOC knowledge assessments	No info	Yes	Post-MOOC knowledge assessment contribute to 100% of grade	Certificate
	27. Lan et al. (2019)	China-Hong Kong	Health	K	Formative; Summative	Computer; self; peer	5; non-graded assessments	No info	Yes; not contribute to grade	Yes	Certificate
	49. Singh & Morch (2018)	Norway	Nature & Environment	K	Formative	Peer	No info	No info	No info	No info	No info
	60. Watson et al. (2017)	USA	Social Science	K	Formative	Self	10	No info	No info	No info	No info
Discussion forum	18. Jia et al. (2019)	China	Medical Health	K	No info	No info	3 case discussions	No info	Yes	10% of grade	No info
	39. Najafi et al. (2017)	Canada	Health, Computer Science, Social Science & Humanities	K; IS	Formative	Peer, instructor	No info	No info	Yes	10% of grade; not contribute to grade	No info
	54. Torres & Beier (2018)	USA	No info	K	No info	Instructor	No info	Yes	Yes	40% of grade	Certificate
Test	18. Jia et al. (2019)	China	Medical Health	K	No info	No info	10 topic tests; in-class test	No info	No info	40% of grade; not contribute to grade	No info
	38. Moreno-Marcos et al. (2018)	Spain	Computer Science	K	No info	No info	5	No info	Yes	75% of grade	No info
	39. Najafi et al. (2017)	Canada	Social Science & Humanities	K; IS	Formative	Computer	2 Web-based tests	Alternative answers	Yes	90% of grade	No info

Table 2.4. An overview of assessment characteristics of cognitive outcomes in the studies reviewed.

Assessment type	Authors (year)	Country/ Region	Discipline	Cognitive Outcomes	Assessment characteristics						
					Purpose	Assessor	Frequency	Feedback	Mandatory	Scoring	Reward
	64. Xing (2019)	USA	Business & Management, Computer Science , Education, the Humanities, Mathematics & Statistics, the Medical Science, the Physical Science, the Professional & Applied Science, interdisciplinary and other fields.	K	No info	No info	No info	No info	No info	No info	No info
Homework	25. Kovanović et al. (2019)	Australia	Computer Science	K; IS	No info	No info	11	No info	Yes	40% of grade	Certificate
	28. Lee (2018)	USA	Engineering	K	Summative	No info	No info	No info	Yes	34% of grade	Certificate
	29. Lee (2019)	USA	Engineering	K	Summative	Computer	No info	No info	Yes	34% of grade	Certificate
Checkpoints	28. Lee (2018)	USA	Engineering	K	Formative	No info	No info	No info	Yes	8% of grade	Certificate
	29. Lee (2019)	USA	Engineering	K	Formative	Computer	No info	No info	Yes	8% of grade	Certificate
Exercise	39. Najafi et al. (2017)	Canada	Computer Science	IS	Formative	Computer	7 exercises in Programming1; 4 exercises in Programming2; 7 exercises in Stats.	No info	Yes	35% of grade in Programming1; 60% in Programming2; 70% of grade in Stats	No info
	60. Watson et al. (2017)	USA	Social Science	K	Formative	No info	No info	No info	No info	No info	No info
User data	51. Sun & Bin (2018)	China	No info	IS	No info	Researcher	1	No	No	No	No
Course project	40. Najafi et al. (2018)	Canada	Computer Science	K	Summative	No info	1 in iOSApp1 and iOSApp2, respectively.	No info	Yes	40% of grade, respectively	No info
Course requirements	1. Almeda et al. (2018)	USA	Humanities	K	No info	Instructor	No info	No info	Yes	No info	Credit, badge
Essay	39. Najafi et al. (2017)	Canada	Social Science & Humanities	IS	Formative	Peer	1 essay in IntroPsych; 1 essay in AboriginalEd; 3 essays in Mental Health	Rubrics	Yes	10% of grade in IntroPsych; 50% of grade in AboriginalEd; 100% of grade in Mental Health.	No info
Labs	25. Kovanović et al. (2019)	Australia	Computer Science	K; IS	No info	No info	7	No info	Yes	60% of grade	Certificate
Online activities	15. Formanek et al. (2017)	USA	Physical Science & Engineering	K	No info	No info	2	No info	Yes	20% of grade	No info
Practical work	22. Khalil & Ebner (2017)	Austria	No info	K	No info	No info	1	No info	No info	No info	Credit, certificate

Table 2.4. An overview of assessment characteristics of cognitive outcomes in the studies reviewed.

Assessment type	Authors (year)	Country/ Region	Discipline	Cognitive Outcomes	Assessment characteristics						
					Purpose	Assessor	Frequency	Feedback	Mandatory	Scoring	Reward
Problem sets	31. Li & Moore (2018)	USA	Chemistry	K	No info	No info	Optional bi-weekly advanced problem sets	No info	No	No info	Certificate
Group embedded figure test	6. Chang et al. (2019)	China-Taiwan	No info	K	No info	Computer	1	No info	No info	No info	No info
Writing projects	8. Comer (2017)	USA	Arts & Humanities	K; IS	Formative	Peer	4	Rubrics	Yes	100% of grade	No info

Note: K = knowledge, IS = intellectual skills.

2.4.2.2. Assessment instruments of behavioral outcomes

In Table 2.5, five types of measurements are shown that assess behavioral outcomes, namely user data, surveys, interviews, observations, and group embedded figure tests, with user data and surveys the two frequently used approaches (e.g., de Lima & Zorrilla, 2017; Khalil & Ebner, 2017; Singh & Mørch, 2018; Z. Wang et al., 2018). The assessors of these instruments were researchers, learner, and computer. Furthermore, no information about several assessment characteristics, including frequency, purpose, feedback, mandatory, scoring, and reward, was reported.

User data recorded the trajectory of learners' engagement and course completion during the learning process (e.g., Li & Baker, 2018; Rabin et al., 2019; Rizvi et al., 2019; Williams et al., 2018; Xing, 2019). Learning analytics techniques and educational data mining are implemented to extract user data to gain insight into learners' behavioral outcomes in MOOCs. For example, Li and Baker (2018) identified learners' behavior based on user data of quizzes submission and videos interaction, and identified subgroups of study patterns (i.e., disengagers, auditors, quiz-takers, and all-rounders). Sunar et al. (2018) categorized learners' behavior into subgroups by the data of learners' contribution and interaction in the discussion threads. Five social patterns were identified: simple, moderately frequent, frequent, and persistent frequent.

The self-report questionnaires were utilized to reveal learners' online learning process. For example, Kizilcec et al. (2017) adopted a five-point scale with 23 statements to investigate learners' engagement with self-regulated learning (SRL) strategies (e.g. elaboration, strategic planning) to support their MOOC learning, and how these strategies drive learners' online behavior. Jung and Lee (2018) modified items from Sun and Rueda (2012), and employed five-point Likert scales to investigate participants' engagement in five courses, namely psychology, design, humanities, musicology, and physics. Additionally, other tools, such as interviews (Krasny et al., 2018), observations (Singh & Mørch, 2018), and group embedded figure tests (Chang et al., 2019) were also used to examine behavioral outcomes of attendees in MOOCs.

Table 2.5. An overview of assessment characteristics of behavioral outcomes in the studies reviewed.

Assessment type	Authors (year)	Country/Region	Discipline	Behavioral Outcomes	Assessment characteristics						
					Assessor	Frequency	Purpose	Feedback	Mandatory	Scoring	Reward
User data	1. Almeda et al. (2018)	USA	Humanities	E	Researcher	1	No	No	No	No	No
	2. Balint et al. (2017)	USA	Engineering	E	Researcher	1	No	No	No	No	No
	3. Bonafini (2017)	USA	Education	E	Researcher	1	No	No	No	No	No
	4. Bonafini et al. (2017)	USA	Physical Science & Engineering	E; COM	Researcher	1	No	No	No	No	No
	7. Chiu & Hew (2018)	China-Hong Kong	No info	E	Researcher	1	No	No	No	No	No
	9. Conijn et al. (2018)	The Netherlands	Computer Science	E	Researcher	1	No	No	No	No	No
	10. de Lima & Zorrilla (2017)	Spain	No info	E	Researcher	1	No	No	No	No	No
	12. Duan et al. (2019)	China	No info	E	Self	1	No	No	No	No	No
	14. Emanuel & Lamb (2017)	USA	Art & Culture	E; CER	Researcher	1	No	No	No	No	No
	16. Harvey et al. (2017)	Australia	Healthcare	E	Researcher	1	No	No	No	No	No
	18. Jia et al. (2019)	China	Medical Health	E; COM	Researcher	1	No	No	No	No	No
	21. Kahan et al. (2017)	Israel	Biology	E; CER	Researcher	1	No	No	No	No	No
	22. Khalil & Ebner (2017)	Austria	No info	E; COM; CER	Researcher	1	No	No	No	No	No
	23. Kizilcec et al. (2017)	USA	Engineering, Computer Science, Management, Transportation, and Education.	E; CER	Researcher	1	No	No	No	No	No
	25. Kovanović et al. (2019)	Australia	Computer Science	E	Researcher	1	No	No	No	No	No
	27. Lan et al. (2019)	China-Hong Kong	Health	E; COM; CER	Researcher	1	No	No	No	No	No
	28. Lee (2018)	USA	Engineering	E; CER	Researcher	1	No	No	No	No	No
	29. Lee (2019)	USA	Engineering	E; COM; CER	Researcher	1	No	No	No	No	No
	30. Li & Baker (2018)	USA	Mathematics	E	Researcher	1	No	No	No	No	No
	35. Luo et al. (2018)	China	No info	E	Researcher	1	No	No	No	No	No
	36. Maldonado-Mahauad et al. (2018)	Chile	Engineering, Education, Management	E	Researcher	1	No	No	No	No	No
	40. Najafi et al. (2018)	Canada	Computer Science	COM	Researcher	1	No	No	No	No	No
	42. Patterson (2018)	USA	No info	E; COM; CER	Researcher	1	No	No	No	No	No
	43. Poquet et al. (2018)	Singapore	Business & Management, Computer Science	E	Researcher	1	No	No	No	No	No

Table 2.5. An overview of assessment characteristics of behavioral outcomes in the studies reviewed.

Assessment type	Authors (year)	Country/Region	Discipline	Behavioral Outcomes	Assessment characteristics						
					Assessor	Frequency	Purpose	Feedback	Mandatory	Scoring	Reward
	45. Rabin et al. (2019)	The Netherlands	No info	E; CER	Researcher	1	No	No	No	No	No
	46. Rizvi et al. (2019)	UK	Science	E	Self	1	No	No	No	No	No
	52. Sunar et al. (2018)	Turkey	Study skills	E; COM	Researcher	1	No	No	No	No	No
	54. Torres & Beier (2018)	USA	No info	E; COM	Researcher	1	No	No	No	No	No
	56. van den Beemt et al. (2018)	The Netherlands	Data Science	E; CER	Researcher	1	No	No	No	No	No
	58. Wang, Anderson & Chen (2018)	China	No info	E	Researcher	1	No	No	No	No	No
	59. Wang, Hu & Zhou (2018)	China	Engineering	E; COM; CER	Researcher	1	No	No	No	No	No
	60. Watson et al. (2017)	USA	Social Science	CER	Researcher	1	No	No	No	No	No
	61. Williams et al. (2018)	USA	Humanities /Liberal Arts, STEM	E	Researcher	1	No	No	No	No	No
	62. Wong et al. (2019)	The Netherlands	No info	E	Researcher	1	No	No	No	No	No
	64. Xing (2019)	USA	Business & Management, Computer Science, Education, Humanities, Mathematics & Statistics, Medical Sciences, Physical Sciences, Professional and Applied Sciences	COM	Researcher	1	No	No	No	No	No
Survey	17. Hudson et al. (2019)	UK	Science, Engineering & Math	E	Self	1	No	No	No	No	No
	19. Jung & Lee (2018)	Republic of Korea	Psychology, Design, Humanities, Musicology, Physics	E	Self	1	No	No	No	No	No
	23. Kizilcec et al. (2017)	USA	Engineering, Computer Science, Management, Transportation, and Education.	E	Self	1	No	No	No	No	No
	33. Li et al. (2018)	China	No info	COM	Self	1	No	No	No	No	No
	37. Meek et al. (2017)	UK	Health & Psychology	E	Self	1	No	No	No	No	No
	44. Qiu & Xu (2018)	USA	Management & Leadership	CER	Self	1	No	No	No	No	No

Table 2.5. An overview of assessment characteristics of behavioral outcomes in the studies reviewed.

Assessment type	Authors (year)	Country/Region	Discipline	Behavioral Outcomes	Assessment characteristics						
					Assessor	Frequency	Purpose	Feedback	Mandatory	Scoring	Reward
	49. Singh & Morch (2018)	Norway	Nature & Environment	E; COM	Self	1	No	No	No	No	No
	50. Stich & Reeves (2017)	USA	No info	COM	Self	1	No	No	No	No	No
	53. Swinnerton et al. (2017)	UK	Health & Psychology	E	Self	1	No	No	No	No	No
	55. van de Oudeweetering & Agirdag (2018)	The Netherlands	Physical Science & Engineering, Social Science, Data Science	COM	Self	1	No	No	No	No	No
	65. Yang & Su (2017)	China-Taiwan	2D Animation Production	E	Self	1	No	No	No	No	No
Interview	34. Loizzo et al. (2017)	USA	No info	COM, CER	Researcher	1	No	No	No	No	No
Observation	49. Singh & Morch (2018)	Norway	Nature & Environment	E; COM	Researcher	1	No	No	No	No	No
Group embedded figure test	6. Chang et al. (2019)	China-Taiwan	No info	E	Computer	1	No	No	No	No	No

Note: E = engagement, COM = course completion, CER = course certificate.

2.4.2.3 Assessment instruments of affective outcomes

In Table 2.6, five types of assessment approaches were utilized to assess the perceptions of learners' various affective outcomes of MOOC learning. Among these approaches, surveys and interviews were the two mostly used, followed by observations, self-assessment, and user data. In the columns of discipline, frequency, and assessor, we see that these instruments were widely used across disciplines. DeBoer et al. (2019), and Kormos and Nijakowska (2017) have conducted surveys at pre-course and post-course, and affective outcomes were evaluated by leaners themselves and researchers. No information is available on the categories of characteristics of feedback, mandatory, scoring, and reward reported.

Likert scales were common applied in self-report questionnaires to reflect learners' perceptions of affective learning gains, such as learning experience, perceived learning benefits, and course satisfaction (e.g., DeBoer et al., 2019; Jung & Lee, 2018; Li, 2019; Rabin et al., 2019; Yang & Su, 2017). For example, Li (2019) applied five-point agreement Likert scales and seven-point likelihood Likert scales for learners to rate the items on course satisfaction and perceived cognitive and affective learning. Rabin et al. (2019) conducted surveys by utilizing 7-point satisfaction and 7-point agreement Likert scales to mirror learners' course satisfaction, perceived course usability, and intention-fulfillment in a nine-week MOOC. In some studies, interviews disclosed the interviewee's evaluation of their learning in MOOCs (e.g., Hudson et al., 2019; Li & Moore, 2018; Watson et al., 2017). For a deeper understanding of group leaders about professional development, leader motivations, and barriers encountered, Krasny et al. (2018) conducted semi-structured interviews with open-ended questions with a small number of students. Other types of measurements, such as observations (Loizzo et al., 2017) and self-assessments (Bonafini, 2017), were employed to reveal learners' evaluation on learning experience and benefits in MOOCs. Additionally, in L. Wang et al. (2018) user data was used to identify learners emotional experiences and emotional states.

Table 2.6. An overview of assessment characteristics of affective outcomes in the studies reviewed.

Assessment type	Authors (year)	Country/Region	Discipline	Affective Outcomes	Assessment characteristics						
					Assessor	Frequency	Purpose	Feedback	Mandatory	Scoring	Reward
Survey	5. Buhr et al. (2019)	Canada	Health	PE	Self	1	No	No	No	No	No
	10. de Lima & Zorrilla (2017)	Spain	No info	SAT	Self	1	No	No	No	No	No
	11. DeBoer et al. (2019)	USA	Neuroscience, Engineering	PB	Self	2	No	No	No	No	No
	17. Hudson et al. (2019)	UK	Science, Engineering & Math	PB	Self	1	No	No	No	No	No
	19. Jung & Lee (2018)	Republic of Korea	Psychology, Design, Humanities, Physics, Musicology,	PB; PE	Self	1	No	No	No	No	No
	20. Jung et al. (2019)	USA	Personal Development	PB; PE	Self	1	No	No	No	No	No
	22. Khalil & Ebner (2017)	Austria	No info	SAT	Self	1	No	No	No	No	No
	24. Kormos & Nijakowska (2017)	UK	Teaching	PB; PE	Self	2	No	No	No	No	No
	25. Kovanović et al. (2019)	Australia	Computer Science	PB; PE	Self	1	No	No	No	No	No
	26. Krasny et al. (2018)	USA	No info	PB; PE	Self	1	No	No	No	No	No
	31. Li & Moore (2018)	USA	Chemistry	PB; PE	Self	1	No	No	No	No	No

Table 2.6. An overview of assessment characteristics of affective outcomes in the studies reviewed.

Assessment type	Authors (year)	Country/Region	Discipline	Affective Outcomes	Assessment characteristics						
					Assessor	Frequency	Purpose	Feedback	Mandatory	Scoring	Reward
	32. Li (2019)	USA	Social sciences, Arts & Humanities, Business, Health, Math & Logic, Physical Science & Engineering, Personal Development	PB; PE; SAT	Self	1	No	No	No	No	No
	33. Li et al. (2018)	China	No info	PB; PE	Self	1	No	No	No	No	No
	37. Meek et al. (2017)	UK	Health & Psychology	PB; PE	Self	1	No	No	No	No	No
	40. Najafi et al. (2018)	Canada	Computer Science	PB	Self	1	No	No	No	No	No
	43. Poquet et al. (2018)	Singapore	Business & Management, Computer Science	PE	Self	1	No	No	No	No	No
	45. Rabin et al. (2019)	The Netherlands	No info	PB; PE; SAT	Self	1	No	No	No	No	No
	47. Rodriguez & Armellini (2017)	Mexico	Study skills	PB	Self	1	No	No	No	No	No
	49. Singh & Morch (2018)	Norway	Nature & Environment	PB; PE	Self	1	No	No	No	No	No
	50. Stich & Reeves (2017)	USA	No info	PB; PE	Self	1	No	No	No	No	No
	53. Swinnerton et al. (2017)	UK	Health & Psychology	PB; PE	Self	1	No	No	No	No	No
	57. Vázquez et al. (2018)	Mexico	Engineering	PB; SAT	Self	1	No	No	No	No	No

Table 2.6. An overview of assessment characteristics of affective outcomes in the studies reviewed.

Assessment type	Authors (year)	Country/Region	Discipline	Affective Outcomes	Assessment characteristics						
					Assessor	Frequency	Purpose	Feedback	Mandatory	Scoring	Reward
	60. Watson et al. (2017)	USA	Social Science	PB; SAT	Self	1	No	No	No	No	No
	65. Yang & Su (2017)	China-Taiwan	Art & Humanities	PB; PE	Self	1	No	No	No	No	No
Interview	17. Hudson et al. (2019)	UK	Science, Engineering & Math	PB	Researcher	1	No	No	No	No	No
	26. Krasny et al. (2018)	USA	No info	PB; PE	Researcher	1	No	No	No	No	No
	31. Li & Moore (2018)	USA	Chemistry	PB; PE	Researcher	1	No	No	No	No	No
	34. Loizzo et al. (2017)	USA	No info	PB; PE	Researcher	1	No	No	No	No	No
	48. Shapiro et al. (2017)	USA	Physical Science & Engineering, Data Science	PB; PE	Researcher	1	No	No	No	No	No
	60. Watson et al. (2017)	USA	Social Sciences	PB	Researcher	1	No	No	No	No	No
	63. Wu et al. (2019)	China-Taiwan	Management & Leadership	PB	Researcher	1	No	No	No	No	No
Observation	34. Loizzo et al. (2017)	USA	No info	PB; PE	Researcher	1	No	No	No	No	No
Self-assessment	3. Bonafini (2017)	USA	Education	PB	Self	1	No	No	No	No	No
User data	8. Comer (2017)	USA	Arts & Humanities	PB	Researcher	1	No	No	No	No	No
	59. Wang, Hu & Zhou (2018)	China	Engineering	PE	Researcher	1	No	No	No	No	No

Note: PE = perceptions of MOOCs experience, PB = perceptions of MOOCs benefits, SAT = course satisfaction.

2.5. Discussion

In this systematic review, we examined (1) what learning outcomes are assessed in MOOCs, (2) what instruments are employed to assess these learning outcomes, and (3) what assessment characteristics these instruments have. Learning outcomes that were measured belong to a wide range of disciplines, the main fields including STEM (i.e., science, technology, engineering, and mathematics), humanities and social science, and medical science. We found that various types of instruments were employed to assess cognitive outcomes, behavioral outcomes, and affective outcomes, and these instruments own their characteristics.

2.5.1. Assessment of cognitive outcomes

Learners' *cognitive outcomes* were measured to mirror the *knowledge* and *intellectual skills* they have obtained and mastered, which were universal outcomes identified in most MOOCs. Twenty types of instruments were utilized to assess these outcomes, such as quizzes, assignments, tests, exams, exercises, and writing projects, in which quizzes, assignments, and exams were frequently used approaches.

In the field of STEM, we found that most of the assessments like quizzes, exams, exercises, tests, labs, homework are built-in the video lectures or follow each course session, with the aim to examine to what extent learners have learned the information that was taught and presented in the subject matter and learning materials. These assessments have different functions that can be used in a formative or summative way. For example, exams are commonly used for a midterm or final summative purposes in the studies reviewed. In another study in a programming course, participants have multiple opportunities for exercise submissions so that they can self-diagnose the problematics and control the learning outcomes themselves, which give the assessment a formative function (Najafi et al., 2017). Generally, the graded assessments are mandatory to participate in, which examine the required learning outcomes and make up the proportion of the final course grades. In most cases, since the courses of STEM are easy to provide standardized and explicit answers in the assessments, the auto-graded selected-response questions and short-answers questions with feedback are quick and efficient for a large number of learners to understand their performance. The elaborate interpretations and feedback of the questions related to the core subject matter could enhance learners' understanding. Moreover, the examinations that focus more on the disciplinary-related and practice-related skills are also part of assessing learning outcomes. For example in (Najafi et al., 2017), the combination of assessment approaches in a Stats MOOC addressed multiple cognitive outcomes: quizzes

examined the content knowledge gains, while complex assignments measured the skills of knowledge application and statistical analysis.

In the MOOCs about humanities and social science, instruments like essays, writing assignments, and writing projects were utilized to evaluate learners' knowledge, critical competencies, and argumentation abilities. We found that peer review and peer feedback with the help of rubrics were features of essay-type assessments. Peer feedback rubrics measured writing-related outcomes, which guided the assessors to check and comment on the components of the writings. The rubrics with structured criteria constructed the road map for both feedback providers and learners towards learning outcomes. For example, in the study of Comer (2017), peer feedback rubrics identified stages of outcomes progress, which learners used for self-reflection and self-critique with the help of peer feedback, which supported them in their writing improvement. Learners indicated that they gained significant benefits in peer reviewing these writings, such as writing proficiency and task understanding. However, one point must be given the attention that students' ability to peer review generally affects to what extent learners perform in essay assignments (Huisman et al., 2018). Explicit interpretations about the utilization of feedback form could perhaps better guide providers and improve the quality of peer feedback Comer (2017). In sum, in MOOCs of humanities and social science, outcome-oriented feedback rubrics can play a scaffolding role in supporting and stimulating learners' performance in essay-type assignments.

Concerning the MOOCs in medical science, examinations, such as exams, tests, and assessments, addressed learners' cognitive learning gains, in which the graded examinations contribute to the proportion of the final grades. The topic tests and in-class tests that were used in Jia et al. (2019), realized formative and summative purpose with both graded and non-graded assessments. The studies reviewed indicated that medical-related courses have components concerning the application of professional knowledge and intellectual skills, with practice-oriented assessments bridging medical theory and authentic practices. Participants learn from case scenarios or patient problems and their performance in the practice cases was assessed. For example, case discussions and virtual patient cases can both support and measure learners' comprehensive knowledge and competencies, such as theoretical and fundamental knowledge, clinic procedures, clinic reasoning, and decision making (Jia et al., 2019; Lan et al., 2019). Teachers led the group discussions to address the questions (Jia et al., 2019), peer and self-assessment measure learners' performance on case problems, and peer solutions contribute to self-reflection (Lan et al., 2019). Furthermore, it is necessary to facilitate learners in higher-level cognitive processing when participating in MOOC forums, which could stimulate learners

to provide more constructive feedback to peers to improve outcomes. This is corroborated by a recent study of Chiu and Hew (2018) that revealed that commenting (higher cognitive processing) on posts in discussion forums have a significant correlation with participants' academic performance. Overall, in medical-related courses, discussion forums could support learners to gain better learning outcomes from medical theory and authentic cases. Instructors, who fulfill the role of leader and facilitator, are essential to engage learners actively in discussion forums.

Furthermore, one could doubt whether the same standard measurement is a valid assessment of the learning outcomes of all learners within one MOOC. In the studies reviewed, the same standard measurement was used to assess all learners in one course, which fails to consider the heterogeneity of learners. Since a diversity of learners participates in MOOCs with various motivations such as course complement, professional needs, satisfying curiosity, connecting with people, their learning needs might differ as well (Zheng et al., 2015). Examples of different learner groups are active learners aim to complete and pass the course, and lurkers do not involve in learning or not complete any assessment tasks at all (Alario-Hoyos et al., 2014). The measurement with same standard within a course might not satisfy all learners' learning needs. One way could be to apply the assessments in each learner group, but differentiated in difficulty and complexity. For example, in the study by Najafi et al. (2018), the higher motivation learner group who were actively engaged in courses achieved significantly better grades than the lower motivation learner group. To take into account the motivation level of these groups, the higher motivation learner group could participate in more difficult and more complex assessment tasks. In contrast, the lower motivation learner group takes the standard assessment tasks or other adaptive assessments. The assessment characteristics of the methods could be helpful to consider appropriate ways to fit learning outcomes. In a word, differentiated learning assessments might align with learners' motivation levels and lead all of them to fulfill learning goals in the end.

2.5.2. Assessment of behavioral and affective outcomes

When we look at the assessment of behavioral and affective outcomes, instruments such as user data, surveys, interviews, and observations were utilized as external evaluation tools to explore different aspects of participants' learning in MOOCs. There are multiple variables of behavioral and affective outcomes that have been explored and different measurements that have been applied.

With regard to assessments of behavioral outcomes, user data was frequently adopted to disclose various sorts of objective learning behavior, which records the trajectory of learners' learning process and captures the moment of what learners did during MOOC learning. Extracting user data can provide insight into how learners behave while learning, such as the times/duration of watching video lectures, the number of posts/replies/comments in discussion forums, and the completion and submission of assessment tasks. These data could provide curriculum teachers quantitative information about the course and reveal to what extent learners are engaged in course learning. For example, the number of videos watched and the number of posts are significantly associated with learners' achievement (Bonafini et al., 2017). Moreover, user data can also help scholars to explore further multiple and complex aspects of the learning process. For example, the different pathways of learners' access to the learning resources and a variety of behaviors can identify what self-regulated learning strategies learners use (Maldonado-Mahauad et al., 2018), what types of social network emerge (de Lima & Zorrilla, 2017), and what problem-solving patterns learners use (Lee, 2019).

Concerning the assessment of affective outcomes, questionnaires were most frequently employed to survey learners' perceptions of learning in MOOCs. The majority of studies have explicitly described the methodology adopted, reported the constructs, reliability and validity of the measurements, and included the items of the scales in detail (e.g., Buhr et al., 2019; DeBoer et al., 2019; Hudson et al., 2019; Kormos & Nijakowska, 2017). Curriculum designers and teachers have access to the complete information and parameters of the survey tools so that they can obtain learners' perceptions in detail by these instruments. However, several studies lacked presenting either the rated items or psychometric properties (e.g., Rodriguez & Armellini, 2017; Valdivia Vázquez et al., 2018), which might impede readers to further understand the results and interpretations of the study. Other instruments like interviews and observations were used for qualitative data collection in the affective domain. Some studies reported methods, such as group discussion, member checking, inter-rater agreement, interviewees review transcripts, which were used to serve for validity checking of the qualitative analysis (e.g., Hudson et al., 2019; Loizzo et al., 2017; Watson et al., 2017; Wu et al., 2019). In sum, questionnaires map learners' perceptions with quantitative methods, and interviews and observations can mirror qualitative views of learners' learning.

This review also found that the cognitive, behavioral, and affective outcomes were correlated. Some of the studies reviewed have gained insight into their correlation. For example, engagement and interaction duration in discussion forums were positive associated with course achievement (Almeda, 2018; Bonafini et al., 2017; Chiu & Hew, 2018), high-quality peer

interactions in peer-reviewed tasks led to a higher performance (Meek et al., 2017), and course completers showed actively social interactions (Sunar et al., 2018). The results of Jung and Lee (2018) indicated that learners' engagement was correlated with learning persistence, teaching presence, perceived usefulness, and academic self-efficacy. Watson et al. (2017) revealed that learners who were rewarded with paid certificates had higher perceived affective learning gains than free certificates receivers, and learners who preferred exploratory activities had a higher sense of course satisfaction than learners in instructor-led activities. These studies have proved that learning outcomes are complex and comprise a variety of variables, and that positive behavioral and affective gains promote the acquisition of cognitive outcomes. The results of these studies might benefit curriculum designers and teachers that they can understand and support learners' learning from multiple perspectives based on the correlation among cognitive, behavioral, and affective learning outcomes, and then optimize the curriculum and instruction in the future.

2.6. Limitations and future research

The main limitation we would like to address here is the possibility of publication bias. Publication bias is a common problem of literature reviews, because studies that show positive or significant results tend to be published more. In our current review study, most of the studies reported the assessment of cognitive outcomes of MOOCs, which might be caused by the fact that most of the positive or significant results are about cognitive outcomes. There might be other studies concerning the assessment of behavioral outcomes and affective outcomes of MOOCs that were not published because these studies did not show significant results. Yet the aim of the current review is descriptive: an exploration of types of learning outcomes and approaches to assess these outcomes. Therefore, this publication bias might not violate the conclusions of our review.

The studies reviewed provide an overview of the assessment of cognitive, behavioral, and affective dimensions of learning outcomes in MOOCs. We hope that this review enables MOOC designers and teachers to have a holistic understanding of diverse learning outcomes and instruments. To better understand learners achieve particular learning outcomes, investigating the motivation of learners to participate in a MOOC and how this changes during a MOOC could be a direction of future research. This could help MOOC designers and teachers to align how learners are motivated, what they want to learn, and what they actually do.

2.7. Concluding remarks

In this review study, learning outcomes were distinguished and various types of assessments in MOOCs were characterized. The current review reveals multiple types of assessment instruments and diverse learning outcomes that extend the previous focus on outcomes like engagement, academic achievement, and attrition (Pilli & Admiraal, 2017) and the findings of the assessment of learning outcomes in reviews of Vera et al. (2017) and Deng et al. (2019). The review findings could benefit curriculum designers and teachers to consider appropriate learning outcomes variables and assessment approaches to serve their assessment needs.

The main outcomes of this review study implicate that:

(1) Considering the assessment of learning outcomes at the beginning of course design, appropriate measurements could support to formulate explicit assessment goals and to instruct learners to work towards learning outcomes.

(2) A combination of knowledge tests and skills tasks can be used to examine cognitive outcomes in a particular MOOC. For example, quizzes and exams can be approaches to examine content knowledge, complex assessment tasks and practical assignments can aim to assess cognitive skills.

(3) Outcome-oriented feedback rubrics are beneficial to support learners' essay performance, and interpretations of the utilization of rubrics could better guide providers to give peer feedback.

(4) A variety of behavioral and affective outcomes reflect multiple aspects of participants learning in MOOCs, which might contribute to teachers better understand and support them to learn.

(5) Assessment tasks throughout the course may differ in difficulty and complexity, which could align with different levels of motivation of the learners.