

Chapter 4

Challenges in Assessments of Soft Skills: Towards Unobtrusive Approaches to Measuring Student Success



Abhinava Barthakur , Vitomir Kovanovic , Srecko Joksimovic ,
and Abelardo Pardo 

Abstract Rapid technological advances, coupled with globalization, have resulted in a changing economy, requiring graduates and students to master not only technical and subject knowledge but also broad, transferable skills for workplace readiness. However, assessing these essential soft skills and competencies beyond the cognitive domain has often relied on questionnaires, surveys and other self-rated scales, which are subjective, often obtrusive in nature, subject to response biases, and lack scalability. In contrast, the pervasive use of educational technology has provided researchers with the opportunity to unobtrusively collect enormous amounts of factual learners' data which has the potential to overcome some of the challenges with questionnaire-based approaches. These unobtrusive measures increase the possibilities of passively evaluating skill acquisition and supporting learners by personalizing learning according to their needs. This chapter outlines a multi-tiered case study and proposes a novel blended methodology, marrying measurement models and learning analytics techniques to mitigate some of these challenges and unobtrusively measure leadership skills in a workplace learning context. Using learners' reflection assessments, several leadership-defining course objectives were quantified, and their progress was assessed over time. The implications of this evidence-based assessment approach, informed by theory, to measure and model soft skills acquisition are further discussed.

Keywords Assessment · Soft skills · Learning analytics · Measurement theory · Online learning

A. Barthakur (✉) · S. Joksimovic · A. Pardo
University of South Australia, Adelaide, SA, Australia
e-mail: abhinava.barthakur@unisa.edu.au; srecko.joksimovic@unisa.edu.au;
abelardo.pardo@unisa.edu.au

V. Kovanovic
Centre for Change and Complexity in Learning, University of South Australia,
Adelaide, SA, Australia
e-mail: vitomir.kovanovic@unisa.edu.au

© The Author(s), under exclusive license to Springer Nature
Switzerland AG 2023

V. Kovanovic et al. (eds.), *Unobtrusive Observations of Learning in Digital Environments*, Advances in Analytics for Learning and Teaching,
https://doi.org/10.1007/978-3-031-30992-2_4

1 Introduction

The changing nature of the modern workplace coupled with technological advances has led to short shelf life for educational practices focused on rote learning within traditional settings, shifting the focus on learners applying their skills and knowledge (Pellegrino, 2017) and on assessment of competencies (Milligan, 2020). Unlike classroom activities that are often constrained to independent learning within disciplinary boundaries, real-world challenges in the modern workplace typically demand collaborative and individual learning approaches to transfer theoretical understandings to practical applications (Ginda et al., 2019). Business leaders, employers, educational stakeholders, and researchers recognize that school success is not the only influential factor determining the economy's success (Kyllonen, 2012) and have called for policies that would support and promote the development of more broad, transferable skills. Organizations deem these broad skills essential for workplace success to solve real-world challenges. Researchers and practitioners lack consensus on the terminology describing these skills (Joksimovic et al., 2020), but they are commonly referred to as 21st century competencies or soft skills (Vockley, 2007). Often differing from context to context, soft skills are well accepted as inherently social and developed through collaboration and networking (Jenkins et al., 2006). Specifically, skills such as communication (oral and written), critical thinking, leadership, problem-solving, and teamwork are some of the most desirable competencies for future graduates (Casner-Lotto & Barrington, 2006; Lai & Viering, 2012).¹

Despite much work being undertaken on promoting soft-transferable skills, they are inherently complex and assessing them is less straightforward (Joksimovic et al., 2020). Martin et al. (2016) note that “even with increasing attention to the importance of 21st century skills, there is still relatively little known about how to measure these sorts of competencies effectively” (*ibid.*, p. 37). The evaluation of the development and acquisition of soft skills is primarily done using introspective approaches such as self-reported questionnaires and inventories (Amagoh, 2009; Ebrahimi & Azmi, 2015), often considered obtrusive. Lai and Viering (2012), and Sondergeld and Johnson (2019), among others, note that the cost-effectiveness, ease of implementation and the ability to provide scores on multiple abilities simultaneously make such approaches popular. However, there are several important challenges associated with these commonly adopted introspective measures, including response biases (Bergner, 2017; Gray & Bergner, 2022), scalability and coverage (Pongpaichet et al., 2022), to name a few. Such challenges make the adoption of such introspective approaches much more challenging, prompting the need to look for alternative approaches.

In contrast, adopting online learning platforms along with educational technologies such as a learning management system provides unobtrusive ways of collecting

¹For convenience and to minimize the multiplicity of terms used to describe the same skills and competencies, we refer to them as soft skills throughout this chapter.

educational data. Relatively recently, educational organizations have started investing in longitudinal learning data, collected at various levels of granularity, to provide insights into teaching quality and understanding the learning process (Joksimovic et al., 2019). The emergence of educational research domains such as *learning analytics* (LA) has demonstrated the potential to support the assessment of learners' skill acquisition through unobtrusive approaches by utilizing fine-grained data collected from educational technologies. Collective sets of LA research in the domain of soft skills measurement have been published in the *Journal of Learning Analytics* across two editions in 2016 (Shum & Crick, 2016) and 2020 (Joksimovic et al., 2020).

In this chapter, we discuss the use of unobtrusive measures to assess soft skills and illustrate a case study on assessing leadership skills. While other skills, such as creativity, critical thinking, and complex problem-solving, received significant attention, there has been little work on assessing leadership skills. In the next section of the chapter, we discuss the various modalities aimed at developing soft skills, some of the challenges associated with the current assessment approaches and, subsequently, the need for more advanced methodologies. To address these challenges, we outline a case study that measured leadership skills across a MOOC study program in three components. In the first component, we rely on an automated machine learning classifier to extract unobtrusive measures from reflective artefacts. We then explore the use of learning analytics techniques and measurement models to evaluate the mastery and acquisition of leadership in a single MOOC. In the final component of the case study, we explore the systematic longitudinal progression of learners developing their skills. We aim to identify some relationships between the learning objectives across multiple courses and how fulfilling the prerequisites in one course helps learners progress in subsequent courses.

2 Background

2.1 *Developing Soft Skills*

While most jobs nowadays demand a broad set of skills to adequately deal with real-world challenges and prepare for an unknown future (Rios et al., 2020), the development of such skills and competencies has been an increasing concern among employers and educators (Shum & Crick, 2016; Haste, 2001). For instance, the Partnership for 21st Century Skills (P21) highlighted higher education's inefficiency in adequately developing these transferable skills (Casner-Lotto & Barrington, 2006), resulting in significant issues with graduate employability. Employers and private organizations encouraged universities and educational institutions to incorporate such skills into their curricula and put greater emphasis on developing complex skills. Additionally, the curricula must be constantly re-evaluated and revised depending on the labour market requirement.

Besides developing soft skills within the traditional classroom settings, workplace learning programs are used to develop soft skills to meet the rapid changes in the modern workforce. Organizations worldwide, for example, are developing professional training courses to deliver the skills and competencies required to tackle the ever-changing work demands (Amagoh, 2009; Burke & Collins, 2005). The need for rapidly changing skill sets and new technological affordances have provided the scope for shifting the focus from classroom learning toward leveraging online settings for developing the necessary workforce skills and professional development within their employees. While some organizations encourage employees to acquire job-relevant skills through off-the-shelf courses, others co-create certification courses along with educational providers to reduce the gap between the skills graduates and employees need to be successful in the modern workforce (Ginda et al., 2019). Moreover, unlike traditional courses in higher education, workplace training usually prioritizes learning processes that focus on transferring content knowledge to practical workplace applications.

Online learning has been increasingly seen as a prominent approach to delivering these workforce programs dedicated to upskilling their employees. These trends have also been accelerated by the recent COVID-19 pandemic, which put online learning at the centre of the educational policy of many governments around the world. One modality of online learning that witnessed growing interest in the domain of professional development is Massive Open Online Courses (MOOCs). Besides providing opportunities for gaining conceptual hold over subject-related knowledge, MOOCs, through their varying pedagogy and self-regulated learning, provide learners with opportunities for developing soft-transferable skills for lifelong learning (Chauhan, 2014). The underlying impact of MOOCs in nurturing these highly valued skills in the labour market allows learners to cultivate knowledge and skills beyond a specific domain. Therefore, their use holds great value from not only developing these skills among learners, but also providing unobtrusive means to collect data.

2.2 Leadership Skills

Numerous skills fall under the umbrella of *soft skills*. Although all these skills are indicative of being effective in dealing with challenges within professional life, Rios et al. (2020) argue that employers do not deem all of them equally essential for their organization. Skills such as written communication, deemed critical for any workplace setting, are missing in 47% of 2-year and 28% of 4-year graduates (Casner-Lotto & Barrington, 2006). In contrast, some 21st century competencies, such as social responsibility, are rarely mentioned in job advertisements (Rios et al., 2020). Therefore, the distinction between the development of novel skills and those adjudged necessary ought to lead educators and policymakers to make educational reforms to decrease learning disparities and improve workforce readiness.

One such essential soft skill that is widely accepted in creating organisational impact and increasingly seen as employment quality is leadership capability (Rohs & Langone, 1997). Leadership skills are considered essential by almost 82% of organizations (Casner-Lotto & Barrington, 2006). Leadership skills contribute to a positive work environment and job satisfaction among employees (Amagoh, 2009). As such, to support the development of leadership skills, various instructional programs are offered in both academic and informal workplace learning settings. Along with providing opportunities to enhance problem-solving capabilities, communication and collaboration, these programs facilitate learning through open-ended and unstructured learning tasks (Joksimovic et al., 2020). Another key aspect of these workplace programs is the emphasis on reflection-promoting activities, encouraging participants to reflect on their learnings and professional experiences (Amagoh, 2009; Burke & Collins, 2005). Such reflective practices show potential in continuously developing skills through purposeful consideration of key concepts and transferring knowledge to real work-life scenarios (Helyer, 2015). Therefore, reflection activities are common educational practices that are used as means to measure the growth and acquisition of skills.

2.3 Challenges of Assessing Soft Skills

The widely adopted P21 framework of 21st century skills have emphasized the need for assessing the learning and acquisition of soft skills to provide formative intervention to steer and support students' performance (Casner-Lotto & Barrington, 2006). Although the various frameworks developed to understand soft skills provide preliminary empirical evidence of their meaning and value (Pellegrino, 2017), unlike measuring "content" or discipline-specific knowledge in classroom settings, assessing soft skills is far more complex and has been of increasing concern for a couple of reasons: there is a lack of coherent understanding of the nature and development of soft skills (Care et al., 2018) and thus, it is hard to quantify them (Joksimovic et al., 2020). Henceforth, researchers have raised several concerns regarding their measurement. Some of the major concerns associated with the assessments of soft skills are as follows:

- *Biases* – Recruiting learners to participate in self-reported scales includes different response biases (Bergner, 2017; Gray & Bergner, 2022). Some of the commonly observed biases are response shift bias (shift in the frame of reference of the measured construct; Barthakur et al., 2022a, b, c; Rohs & Langone, 1997), social desirability bias (rejecting undesirable characteristics and faking socially desirable traits; Nederhof, 1985), biases that result from participants resorting to extreme ends of Likert scale (Bachman & O'Malley, 1984), among others. As such, although it is assumed that participants are honest while answering these surveys and questionnaires, they are replete with biases that cannot be ignored.

- *Scalability* – The time-consuming, costly, and labour-intensive aspect of incorporating self-reported scales as means of assessing soft skills limits the frequency and coverage of these approaches. Self-reported measures lack scalability and cannot be deployed to measure skill development among a wider audience. Similarly, the administration of survey-type questionnaires does not guarantee total participation (Pongpaichet et al., 2022). Furthermore, monitoring the progression of these complex skills over time is vital and critical for enhancing learning outcomes (Dawson & Siemens, 2014). However, administering the same questionnaire repeatedly to measure growth can result in burnout (Sutherland et al., 2013).
- *Pre/post-test* – While adopting a pre- and post-test approach to measure skills development has been a prominent approach, such techniques do not account for the learning taking place during the study period. Pre-post assessment models developed for measuring leadership skills are usually deployed before and after learning content delivery (Amagoh, 2009) and provide snapshots of learning overtime. As such, they cannot capture the learning progression of the learners through the different stages of skills development and how their learning is associated with the development.
- *Active assessment* – Learners are required to participate in assessment questionnaires and surveys during the study period; thus, interfering obtrusively with their learning processes. As such, there is a need to adopt unobtrusive methods to quietly assess soft skills and allow instructors to monitor learners' development and growth without interrupting the study flow (Pongpaichet et al., 2022).
- *Analytical techniques* – Traditional measurement models used in the field of psychometrics and learning assessments do not utilize the educational data generated by online learning platforms to the full extent. Measurement models used by psychometricians usually rely on fixed-item responses by participants to measure learners' knowledge about subject content without necessarily considering the learning strategies adopted by participants while solving tasks. Traditional assessment techniques developed for the analysis of test responses cannot be applied to educational trace data. As such, the existing approaches do not consider the learning process and what learners do and only focus on the learning outcome. In this regard, there is a need to link the assessment of student learning outcomes and their learning behaviour and strategies to effectively identify the overall progress. In contrast, the fields of Educational Data Mining (EDM) and Learning Analytics (LA) have utilized trace data to provide unobtrusive means of assessing the learning strategies adopted by learners within MOOCs; thus, contributing to a richer understanding of the complex behaviour associated with student learning (Dawson & Siemens, 2014) without interfering with the dynamic learning process. Also, the use of trace data collected from various educational technologies eliminates biases generated from self-reported measures and the effort of administering additional instruments to collect data (Gray & Bergner, 2022). However, the statistical relations found in these LA studies only demonstrate that these patterns are unlikely random but can be inconsequential in judging an individual's learning (Milligan, 2020). The probabilistic dependency of

the observed variables on the targeted latent skill/learning objectives is often missing within LA (Mislevy et al., 2012). Although there are individual limitations in both these educational assessment fields, several studies have adopted multi-disciplinary techniques that draw on the strengths of one another for providing a holistic assessment of soft skills (Milligan & Griffin, 2016).

While the persistence of these challenges limits the measurement of soft skills, more reliable measures can be achieved through the careful consideration of unobtrusive approaches that go beyond self-reported scores. Therefore, by building on some of the earlier works in LA and implementing advanced analytical methods intersecting measurement models, we propose more scalable and unobtrusive means of assessing soft skills.

3 Case Study

3.1 Study Context

This chapter extracts data from an online professional learning program to develop leadership capabilities among the employees of a large global US corporation. The participants of the program were full-time working professionals and were mainly from the engineering and management domains, with varying professional backgrounds ranging from fresh graduates to individuals with over 15 years of experience. Delivered as a part of workplace training, this program was hosted in the Open_edX platform and was made available for free to all its employees.

This program consisted of a series of four Massive Open Online Courses (MOOCs) covering different aspects of leadership development. The first course of the program was scheduled for 4 weeks, while the remaining three courses were 3 weeks long each, delivered consecutively with a week-long break between two MOOCs. Also, these were asynchronous MOOCs that were designed to deliver several leadership learning objectives through recorded learning videos and related learning modules. Additionally, the MOOCs also included various formative assessments such as quizzes and self-reflection questions and summative essay assessments on several leadership concepts. All these assessments contributed to the certification grade for each MOOC. The second course of the program, however, followed a different instructional design and was left out of the analyses. The three components (MOOCs) of leadership investigated were – *understanding organizational strategy and capability, leading change in organizations, and discovering and implementing individual leadership strengths*.

In this chapter, we are particularly interested in the assessment of the self-reflection answers and use it as a proxy to comprehend and quantify leadership development (Helyer, 2015). The other kinds of formative assessments, such as the polls and the multiple-choice questions, allowed multiple attempts and prompted learners with hints. As such, the answers to these formative assessments may not adequately measure the development of the skills (Barthakur et al., 2022a).

Self-Reflection

Bookmark this page

APPLY ★



Reflection Question

0.0/10.0 points (graded)

Can you think of a time that you would have had a better experience or outcome if you had known about strategy? Is there an opportunity, by becoming more familiar with strategy, to align your efforts with it?

Write 3-5 sentences on what you experienced and how it would go better if you knew the strategy framework you work within.

Fig. 4.1 An example of a self-reflection assessment question

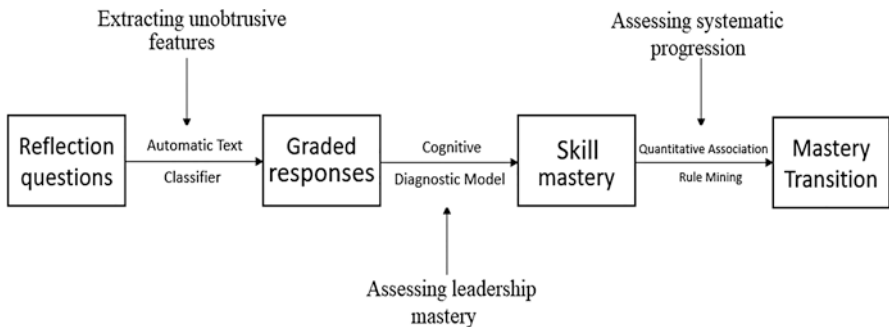


Fig. 4.2 The methodological pipeline used in the outlined case study. (Adopted from Barthakur et al., 2022a)

The self-reflection questions used within these MOOCs were content-specific in the sense that learners were encouraged to reflect on their learnings and experiences from leadership perspectives (Fig. 4.1). While discourse analysis and, more particularly, automated assessment of reflection has been studied by LA researchers for some time (Buckingham Shum et al., 2017; Jung & Wise, 2020; Ullmann, 2019), there is limited research on the assessment of reflection depth by adult learners (Barthakur et al., 2022b). Furthermore, although literature shows the role of reflection in skill development (Densten & Gray, 2001; Helyer, 2015; Wu & Crocco, 2019), there is a dearth of studies focusing on using reflection assessments as an unobtrusive means for evaluating soft skills.

The overview of the methodological pipeline adopted in the case study is provided in Fig. 4.2. In this, we adopt a blended methodology intersecting LA techniques and psychometric measurement models (Drachler & Goldhammer, 2020).

This novel methodology draws on the strengths of different disciplines and mitigates the challenges listed in Sect. 2.3, thus, providing a means for unobtrusively measuring leadership skills.

3.2 *Extracting Unobtrusive Measures*

As we previously discussed, assessments evaluating soft skills are often actively administered, requiring learners to respond explicitly to questionnaires or other self-reported scales (Pongpaichet et al., 2022). In contrast, the analysis of the text responses to self-reflection questions can provide an unobtrusive means of evaluating leadership growth and acquisition, provided that researchers can extract features indicative of leadership skill mastery.

In one of our recent works (Barthakur et al., 2022b), we outlined a methodology to extract unobtrusive features from written reflective artefacts by implementing quantitative content analysis (Krippendorff, 2003) and developed an automated assessment system. The reflection responses varied in length and in the range of 17–393 words, with an average of 74 words utilized across each of the fifteen different questions. Extracting data from 771 out of the 861 learners who attempted all the reflection questions, the responses were categorized into four different hierarchical levels depending on the depth of reflection exhibited. These four levels were coded in accordance with a reflection framework developed by Kember et al. (2008) and are as follows – No-reflection, Understanding, Simple reflection and Critical reflection.

Two independent human coders manually graded a hundred answers each for the first four questions. When inter-rater reliability of 0.70 was achieved, the workload was equally divided to code the remaining answers to the same four questions. Using the manually coded responses as the training set and extracting several linguistic features from the written artefacts (such as Linguistic Word Count Inquiry, Coh-Metrix, n-grams, and readability index, among others), a machine learning classifier was trained to automatically analyse the answers to the remaining reflective assessments from the first MOOC. The performance of these models was judged based on their accuracy (closeness of the predicted values to the true values) and AUC ROC (Area Under the Curve – Receiver Operating Characteristics) values. While the model achieved a moderate accuracy of 0.66 and an AUC ROC of 0.88, the focus of the study was on establishing explainable insights rather than achieving higher accuracy through the implementation of AI black boxes (Dawson et al., 2019; Sartori & Theodorou, 2022).

Overall, the use of an automatic assessment approach provided the means for extracting unobtrusive measures of four reflection levels that were used to build models assessing the mastery of leadership learning objectives. Besides categorizing the responses into different levels based on the depth of reflection, the top twenty linguistic features predictive of the four levels were also analysed (Fig. 4.3). These

Feature	Description	SHAP	Self-Reflection Levels			
			No-Reflection	Understanding	Simple Reflection	Critical Reflection
WC	Word count of each answer	109.07	46.86 (22.91)	64.73 (22.58)	104.32 (33.82)	164.85 (52.27)
cm.WRDPRO	Word information: Pronoun indices	29.33	10.13 (5.06)	15.23 (5.98)	24.22 (9.26)	37.76 (14.63)
cm.WRDPRPis	Word information: First person singular pronoun incidence	28.6	11.12 (5.73)	14.82 (5.61)	23.72 (7.82)	37.3 (11.14)
cm.LDVOCda	Lexical Diversity: Lexical diversity, VOCd, all words	17.75	0.39 (0.41)	0.69 (0.28)	0.81 (0.05)	0.82 (0.03)
MD (POS Tagging)	Modal Verbs (e.g., can, could)	12.67	0.57 (0.95)	0.78 (1.02)	1.32 (1.27)	2.24 (1.85)
Product	Frequency of "Product" Unigram	11.58	0.18 (0.52)	0.57 (1.0)	0.46 (1.05)	0.7 (1.51)
Thing	Frequency of "Thing" Unigram	6.19	0.2 (0.55)	0.13 (0.45)	0.23 (0.64)	0.37 (0.88)
cm.DRPVAL	Syntactic pattern density: Agentless passive voice density incidence	5.59	11.96 (6.02)	16.82 (6.01)	26.54 (8.89)	41.32 (13.48)
NN (POS tagging)	Singular nouns (e.g., business, college)	5.16	7.86 (4.25)	11.4 (5.02)	18.0 (7.46)	28.23 (10.1)
Resource	Frequency of "Resource" Unigram	4.73	0.13 (0.37)	0.37 (0.71)	0.27 (0.7)	0.46 (1.03)
IN (POS tagging)	Preposition or subordinating conjunction	4.03	0.67 (0.86)	1.1 (1.05)	1.84 (1.45)	3.09 (2.01)
cm.CRFANPa	Referential Cohesion: Anaphor overlap, all sentences	3.61	0.25 (0.38)	0.37 (0.36)	0.41 (0.33)	0.43 (0.31)
Control	Frequency of "Control" Unigram	3.4	0.19 (0.6)	0.12 (0.54)	0.2 (0.71)	0.29 (0.87)
Would	Frequency of "Would" Unigram	3.18	0.38 (0.74)	0.51 (0.83)	0.89 (1.08)	1.52 (1.46)
Licw.Dic	Dictionary words	2.98	88.67 (6.54)	88.48 (5.54)	87.52 (5.0)	88.1 (4.42)
cm.DRNEG	Syntactic pattern density: Negation density, incidence	2.91	8.59 (4.33)	11.3 (4.19)	18.17 (6.17)	28.34 (8.67)
difficult_words	Words not matching the Dale-Chall list of "familiar" words	2.44	10.21 (5.03)	13.93 (5.55)	21.81 (8.26)	33.17 (12.07)
Licw:focuspresent	Time orientation: focus towards present	2.4	10.54 (5.1)	8.88 (4.57)	8.09 (4.07)	8.51 (3.67)
Licw.money	Personal concerns: Money (e.g., audit, cash, owe)	2.05	0.77 (1.8)	1.2 (1.9)	1.07 (1.58)	0.88 (1.14)
Licw.you	Personal pronouns: Second person (e.g., you, your)	1.85	0.79 (2.12)	0.72 (2.11)	0.25 (0.96)	0.34 (0.91)
Licw.tentat	Cognitive processes: Tentative (e.g., maybe, perhaps)	1.82	2.27 (2.58)	2.3 (2.27)	2.15 (1.83)	2.39 (1.68)
Project	Frequency of "Project" Unigram	1.81	0.2 (0.49)	0.61 (0.99)	0.63 (1.23)	0.85 (1.52)
PRP (POS Tagging)	Personal pronouns	1.73	1.69 (1.6)	2.09 (1.93)	3.51 (2.29)	5.07 (3.3)
Licw.insight	Cognitive Processes: Insight (e.g., think, know, consider)	1.67	4.16 (3.5)	3.47 (2.92)	3.7 (2.49)	3.54 (2.02)
Licw.Ner	Negations (e.g., no, not, never)	1.64	0.49 (0.96)	0.71 (1.18)	1.45 (1.7)	2.32 (2.38)
Licw.Adverb	Adverbs (e.g., really, very)	1.58	3.29 (3.01)	3.1 (2.51)	3.04 (1.93)	3.25 (1.6)
cm.SYNMEDpos	Syntactic Complexity: Minimal Edit Distance, part of speech	1.37	8.26 (3.67)	10.43 (3.48)	15.52 (4.57)	21.69 (5.31)
RBR (POS tagging)	Adverbs, comparative	1.2	1.52 (1.58)	1.89 (1.66)	3.04 (2.16)	5.21 (2.92)
cm.SYNMEDlem	Syntactic Complexity: Minimal Edit Distance, lemmas	1.19	4.85 (5.05)	5.71 (4.48)	5.73 (3.77)	7.45 (4.3)
People	Frequency of "People" Unigram	1.15	0.07 (0.28)	0.05 (0.24)	0.1 (0.38)	0.14 (0.43)

Fig. 4.3 Top twenty features summary and their association with the four reflection levels. (Adopted from Barthakur et al., 2022b)

features are ranked based on their SHAP score (a unified measure of feature importance), and the association with the four levels is also provided.

While some of the findings echo that of previous studies, such as higher word count being indicative of a higher level of reflective practice, several newer insights regarding reflection in relation to skill development were discovered. For instance, it was observed that learners tend to describe more about their present learning and professional experiences while developing skills compared to other learners in more traditional settings focusing on past events (Kovanovic et al., 2018). Another observation, captured through the readability index, includes the use of more complicated phrases in higher levels of reflective text, while learners engaging in shallow reflection are less expressive (with fewer word counts) and tend to rely on simple dictionary words. Also, the use of first-person and second-person (such as you and your) personal pronouns can help identify the depth of reflection exhibited by the learners. Such insights were previously unknown, are critical for comprehending (leadership) skill development, and have important practical implications (Barthakur et al., 2022b). Based on the reflection levels, such findings can also provide opportunities for supporting and scaffolding learners. Learners demonstrating shallow reflection can be provided real-time feedback and enhance their skill acquisition during the learning process.

3.3 *Assessing Leadership Mastery*

In Sect. 2.2, we highlighted the role of reflective practices in different educational contexts and, more particularly, in developing leadership skills. Extending the above methodology and the extracted unobtrusive measures of reflection on leadership concepts, another research project evaluated the mastery of leadership skills defined as learning objectives in MOOCs (Barthakur et al., 2022c). While the limitation of lack of probabilistic dependencies between the observed variables and the latent learning constructs is often discussed in LA studies (Milligan, 2020), in this second component of the case study, mastery of (five) latent leadership objectives was calculated using the ordinal four-level graded reflections and a probabilistic relationship was established.

In this example, we divided the analysis into two steps – providing an assessment of the mastery of the individual skills based on the reflection grades and finding clusters of students based on their mastery of all skills in the course. First, a measurement model (cognitive diagnostic model, CDM; Lee & Sawaki, 2009; Rupp et al., 2010) was implemented using the four-level graded reflection responses as the input. CDMs are person-centred models that are used when empirical information about latent skills and attributes is sought (Rupp & Templin, 2008). CDMs in this case study were used to calculate probabilities of (latent) leadership skill mastery for all the learners based on their written artefacts. These models provide information about the extent of mastery of these latent skills, in the range of zero to one. A probability closer to one demonstrates higher mastery, while probabilities on the

other side of the spectrum closer to zero indicate lower levels of mastery. Out of the several types of CDM models, a generalized model was chosen given its generalizability and relaxed nature (devoid of any strong conjunctive or disjunctive assumptions). Usually, most CDM models are constrained and require fulfilling several assumptions as compared to the generalized CDM. Furthermore, due to the ordinal nature of the graded reflection data, a sequential CDM model was used for this analysis (Barthakur et al., 2022c). In the second and final step of the analysis, the learners were then categorized into different groups using a clustering algorithm to develop a holistic understanding of skill mastery at the cohort level in the entire MOOC (Fig. 4.2).

As mentioned earlier, these leadership skills within each MOOC are defined as learning objectives. The probabilities calculated by the CDM models represent the extent of mastery for the various individual learning objectives; probabilities closer to one generally indicate higher mastery. Such results provide diagnostic information about their acquisition and mastery across several leadership components inferred from the reflective responses. Additionally, based on the probabilities of skill mastery, the clustering algorithm identified four distinct learning profiles (Fig. 4.4). These profiles were labelled depending on the average learning objective mastery. It was observed that the latter learning objectives of the MOOC had a higher mastery rate while the lowest mastery across the first learning objective across all the profiles. We supported these findings through leadership theory and propose that the learners were building on the contents associated with the earlier learning objectives to exhibit superior mastery as the course progressed. The gradual shift in the learning objective mastery highlight “the effect of the design and sequencing of individual learning objectives on content mastery” (Barthakur et al., 2022a, b, c, p. 17).

This component of the case study extends the discussion of understanding and assessing reflection answers by operationalizing meaningful latent learning objectives. Based on the depth of reflection exhibited by the learners in the written artefacts, a measurement model was implemented to calculate the mastery of various leadership learning objectives. An important implication of such an analysis is that it shifts the focus from evaluating learners’ cognitive knowledge based on their final course grades to the assessments of skill mastery. It also advances the discussion of learner profiling by categorizing learners based on the evidenced learning objective mastery, which was previously done using behavioural engagement data and final course grades. Moreover, while the study in Sect. 3.2 can be used for scaffolding learners with individual reflection questions, the findings from this component allow instructors and other stakeholders to provide pedagogical interventions depending on learners’ mastery of learning objectives and support learners with specific content within the course. From the methodological standpoint, this work provides a novel blended methodology approach marrying learning analytics and measurement theory models to measure soft skills. The two studies discussed above combined provide an overview of the novel implementation of unobtrusive approaches to soft skills assessment in digital learning settings, opening avenues to

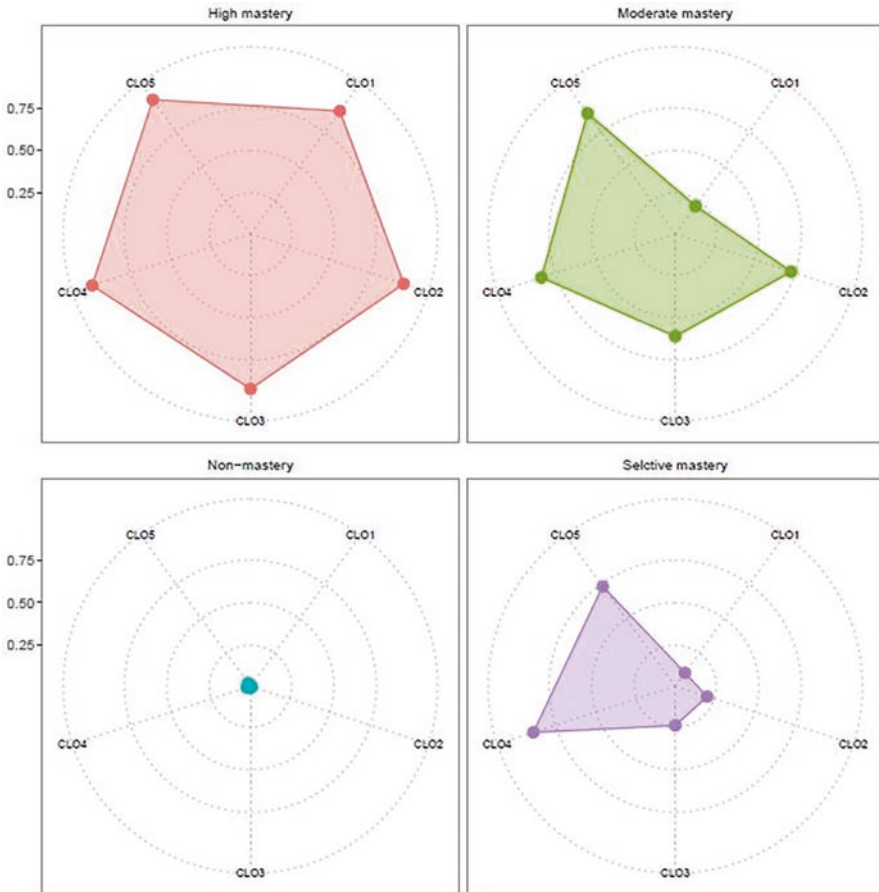


Fig. 4.4 Learner profiles are based on average learning objective mastery. (Adopted from Barthakur et al., 2022c)

extend the methodology to measure their longitudinal progress and growth over time and across several courses.

3.4 Assessing Systematic Progression

Effective and thoughtful sequencing of courses and learning contents are critical for allowing learners to successfully navigate and acquire knowledge while traversing through a study program (Dawson & Hubball, 2014). Study programs with either flexible or restricted pathways, when effectively structured, can often reduce learners' cognitive overload and enhance academic performance (Barthakur et al., 2022a). While the effectiveness of the courses and learning content sequencing are

primarily evaluated using introspective peer-review approaches, these measures have several drawbacks, as suggested in Sect. 2.2. However, the introduction of online micro-credential programs, such as the one discussed in this case study, opens avenues to collect trace data to evaluate the systematic progression of learners across multiple courses.

This work extracted data from 771 learners who engaged with the self-reflection assessments in at least two courses, allowing us to analyse the transitions and understand the pre- and post-requisites of the courses. Building on the previous two research projects, in the final component of the case study, we explore the relationship between several learning objectives across three different MOOCs of the leadership development study program. In the works of Barthakur et al. (2022a), a three-step blended methodology was outlined to automatically evaluate these relationships based on the empirical assessment data across the whole MOOC study program (*ibid*). More specifically, using the machine learning classifier described above to automatically grade the reflective artefacts, the mastery of learning objectives was calculated using multiple CDM models across the entire study program.

In the third stage of the methodological pipeline, a Quantitative Association Rule Mining (QARM; Salleb-Aouissi et al., 2007) was implemented to investigate learners' transitions in learning objective mastery when traversing across the MOOC program. QARM is similar to general association rules with the exception of numerical attributes involved on either side of the rule. For instance, while a general association rule can be expressed as {Butter} → {Milk, Flour}, quantitative association rules are more advanced and can be expressed as {2 Butter} → {3 Milk, 1 Flour}. In this current example, the probabilities calculated from the CDMs in the previous step were converted into three ordinal levels – low mastery (probabilities below 0.60), medium mastery (between 0.60 and 0.80) and high mastery (above 0.80). In doing so, the ordinal levels serve as adequate input to the QARM algorithm and support the identification of the learning objective mastery relationship.

From the first part of the analysis, it was observed that the learners exhibited varying probabilities of learning objective mastery across the three courses. However, the unique contribution of this example is the analysis of mastery transition and understanding how prerequisites in a course affect the mastery of content in the subsequent courses of a study program. Barthakur et al. (2022a) traced some of these findings in various seminal (Quinn, 1988) and modern (Corbett, 2021; Corbett & Spinello, 2020) leadership theories and frameworks. Interpreting the mastery transitions (Fig. 4.5), it was observed that higher mastery across the five leadership objectives in the first MOOC resulted in higher mastery across the first (3.1) and third (3.3) objectives of the third MOOC. On the contrary, failing to demonstrate high command over the learning objectives of the first MOOC can significantly affect the mastery of the last learning objective (3.4) of the third MOOC. Similar observations of low mastery can be made in the second objective (4.2) of the fourth MOOC when failing to master the leadership objectives of the third MOOC on *leading change in organizations*. Such a relationship echoes Quinn's (1988) theories of the role of effective leadership in facilitating change to enhance organizational performance.

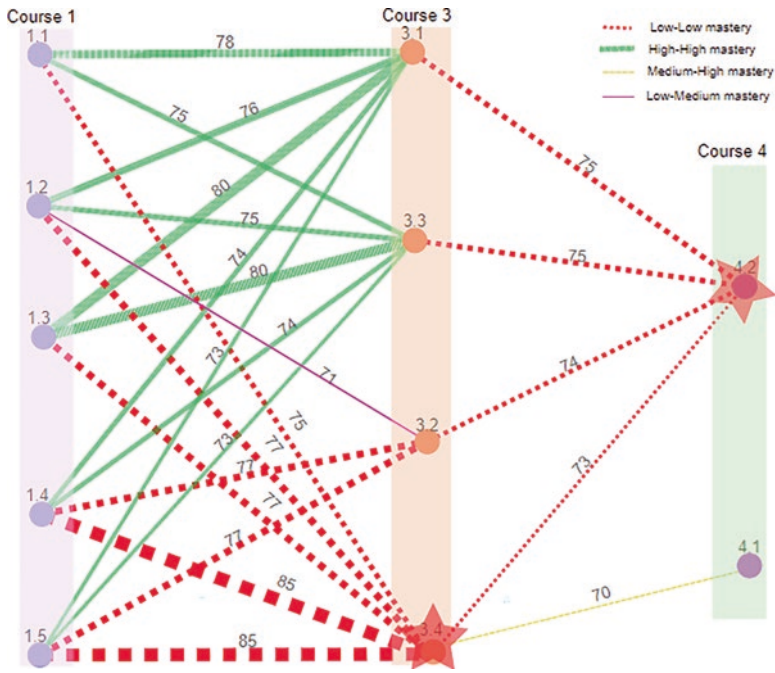


Fig. 4.5 Transitions in leadership objective mastery across a study program. (Adopted from Barthakur et al., 2022a)

Using an evidence-based approach, this work contributes to our understanding of learners’ skill mastery within and across multiple MOOCs in the study program and how they transition over time. Such findings can provide instructors with information about students’ learning which in turn can be used to provide pedagogically informed decisions. For instance, learners exhibiting lower mastery in the first course can be supported with additional resources to successfully complete the final objective (3.4) of the third MOOC. Similarly, these findings can be used for gathering diagnostic fine-grained information about the mastery of individual learning objectives that go beyond analysing learners’ success based on final course grades. Finally, from the perspective of the course designers, instructors, and researchers, this will allow for investigating the ordering of learning objectives and courses to reduce cognitive overload and enhance student learning experiences.

4 Conclusion

The importance of soft skills in the modern workforce has been extensively discussed in the last few decades. Several frameworks have been conceptualized to comprehend and promote the development of these complex skills (Casner-Lotto &

Barrington, 2006). While significant efforts were made in terms of promoting *soft skills*, there were significant challenges in the way these competencies and skills were measured. Most previous research has measured skill development through the use of subjective questionnaire-type measures. However, these measures are often associated with several biases and cannot guarantee total participation. The scalability of such approaches is also questionable.

In this chapter, we illustrate some of these challenges that are associated with the current practices of soft skill assessment and a need for evidence-based approaches for measuring soft skills. A case study, divided into three components (Fig. 4.2), is outlined, describing a data-driven methodology using unobtrusive features for measuring leadership skill mastery and acquisition. Extracting unobtrusive features from learners' self-reflection artefacts in a MOOC study program, responses were automatically graded, and leadership mastery was calculated using a measurement model. The interdependencies of the skills' mastery were further analysed to study the transitions over time. Such a methodology can be easily extended to extract several other unobtrusive features (any hierarchically graded assessments, such as in the form of correct, incorrect, and partially correct responses) from digital learning environments to assess different soft skills.

The underlying premise of the work presented in this chapter revolves around advancing research related to the assessment of complex soft skills. This chapter outlines three studies that illustrate a blended methodology by combining learning analytics and psychometrics to measure leadership skills by collecting digital assessment data from a professional development MOOC program. These unobtrusive approaches to data extraction provide the opportunity to passively measure the development of skills without interfering with the learning processes. The assessment models discussed are fully automatic and thus have the potential to be implemented at scale. Finally, the generalizability of the approach allows the assessment of other skills in varying contexts.

References

- Amagoh, F. (2009). Leadership development and leadership effectiveness. *Management Decision*, 47(6), 989–999. <https://doi.org/10.1108/00251740910966695>
- Bachman, J. G., & O'Malley, P. M. (1984). Yea-saying, Nay-saying, and going to extremes: Black-White differences in response styles. *Public Opinion Quarterly*, 48(2), 491–509. <https://doi.org/10.1086/268845>
- Barthakur, A., Joksimovic, S., Kovanovic, V., Corbett, F. C., Richey, M., & Pardo, A. (2022a). Assessing the sequencing of learning objectives in a study program using evidence-based practice. *Assessment & Evaluation in Higher Education*, 1–15. <https://doi.org/10.1080/02602938.2022.2064971>
- Barthakur, A., Joksimovic, S., Kovanovic, V., Ferreira Mello, R., Taylor, M., Richey, M., & Pardo, A. (2022b). Understanding depth of reflective writing in workplace learning assessments using machine learning classification. *IEEE Transactions on Learning Technologies*, 1. <https://doi.org/10.1109/TLT.2022.3162546>

- Barthakur, A., Kovanovic, V., Joksimovic, S., Zhang, Z., Richey, M., & Pardo, A. (2022c). Measuring leadership development in workplace learning using automated assessments: Learning analytics and measurement theory approach. *British Journal of Educational Technology*. <https://doi.org/10.1111/bjet.13218>
- Bergner, Y. (2017). Measurement and its uses in learning analytics. In C. Lang, G. Siemens, A. Wise, & D. Gasevic (Eds.), *Handbook of learning analytics* (1st ed., pp. 35–48). Society for Learning Analytics Research (SoLAR). <https://doi.org/10.18608/hla17.003>
- Buckingham Shum, S., Sándor, Á., Goldsmith, R., Bass, R., & McWilliams, M. (2017). Towards reflective writing analytics: Rationale, methodology and preliminary results. *Journal of Learning Analytics*, 4(1), 10.18608/jla.2017.41.5.
- Burke, V., & Collins, D. (2005). Optimising the effects of leadership development programmes: A framework for analysing the learning and transfer of leadership skills. *Management Decision*, 43(7/8), 975–987. <https://doi.org/10.1108/00251740510609974>
- Care, E., Griffin, P., & Wilson, M. (Eds.). (2018). *Assessment and teaching of 21st century skills: Research and applications*. Springer. <https://doi.org/10.1007/978-3-319-65368-6>
- Casner-Lotto, J., & Barrington, L. (2006). Are they really ready to work? Employers' perspectives on the basic knowledge and applied skills of new entrants to the 21st century U.S. workforce. In *Partnership for 21st century skills*. Partnership for 21st Century Skills. <https://eric.ed.gov/?id=ED519465>
- Chauhan, A. (2014). *Massive open online courses (MOOCs): Emerging trends in assessment and accreditation* (p. 12).
- Corbett, F. (2021). *Emergence of the connectivist leadership paradigm: A grounded theory study in the Asia region*. Theses and dissertations. <https://digitalcommons.pepperdine.edu/etd/1194>
- Corbett, F., & Spinello, E. (2020). Connectivism and leadership: Harnessing a learning theory for the digital age to redefine leadership in the twenty-first century. *Heliyon*, 6(1), e03250. <https://doi.org/10.1016/j.heliyon.2020.e03250>
- Dawson, S., & Hubball, H. (2014). Curriculum analytics: Application of social network analysis for improving strategic curriculum decision-making in a research – Intensive university. *Learning Inquiry*, 2(2), 59–74.
- Dawson, S., & Siemens, G. (2014). Analytics to literacies: The development of a learning analytics framework for multiliteracies assessment. *International Review of Research in Open and Distance Learning*, 15, 284–305. <https://doi.org/10.19173/irrodl.v15i4.1878>
- Dawson, S., Joksimovic, S., Poquet, O., & Siemens, G. (2019). Increasing the impact of learning analytics. In *Proceedings of the 9th international conference on learning analytics & knowledge* (pp. 446–455). <https://doi.org/10.1145/3303772.3303784>
- Densten, I., & Gray, J. (2001). *Leadership development and reflection: What is the connection?* (p. 15). [http://lst-iiep.iiep-unesco.org/cgi-bin/wwwi32.exe/\[in=epidoc1.in\]/?T2000=013168/\(100\)](http://lst-iiep.iiep-unesco.org/cgi-bin/wwwi32.exe/[in=epidoc1.in]/?T2000=013168/(100)). <https://doi.org/10.1108/09513540110384466>
- Drachsler, H., & Goldhammer, F. (2020). Learning analytics and eAssessment – Towards computational psychometrics by combining psychometrics with learning analytics. In D. Burgos (Ed.), *Radical solutions and learning analytics: Personalised learning and teaching through big data* (pp. 67–80). Springer. https://doi.org/10.1007/978-981-15-4526-9_5
- Ebrahimi, M. S., & Azmi, M. N. (2015). New approach to leadership skills development (developing a model and measure). *Journal of Management Development*, 34(7), 821–853. <https://doi.org/10.1108/JMD-03-2013-0046>
- Ginda, M., Richey, M. C., Cousino, M., & Börner, K. (2019). Visualizing learner engagement, performance, and trajectories to evaluate and optimize online course design. *PLoS ONE*, 14(5), e0215964. <https://doi.org/10.1371/journal.pone.0215964>
- Gray, G., & Bergner, Y. (2022). A practitioner's guide to measurement in learning analytics – Decisions, opportunities, and challenges. In *Handbook of learning analytics* (2nd ed., pp. 20–28).
- Haste, H. (2001). Ambiguity, autonomy and agency: Psychological challenges to new competence. *Defining and Selecting Key Competencies*, 93–120.

- Helyer, R. (2015). Learning through reflection: The critical role of reflection in work-based learning (WBL). *Journal of Work-Applied Management*, 7(1), 15–27. <https://doi.org/10.1108/JWAM-10-2015-003>
- Jenkins, H., Clinton, K., Purushotma, R., Robison, A. J., & Weigel, M. (2006). *Confronting the challenges of participatory culture: Media education for the 21st century*. MacArthur Foundation.
- Joksimovic, S., Kovanovic, V., & Dawson, S. (2019). The journey of learning analytics. *HERDSA Review of Higher Education*, 6, 37–63.
- Joksimovic, S., Siemens, G., Wang, Y. E., San Pedro, M. O. Z., & Way, J. (2020). Editorial: Beyond cognitive ability. *Journal of Learning Analytics*, 7(1), 1–4. <https://doi.org/10.18608/jla.2020.71.1>
- Jung, Y., & Wise, A. F. (2020). How and how well do students reflect? Multi-dimensional automated reflection assessment in health professions education. In *Proceedings of the tenth international conference on learning analytics & knowledge* (pp. 595–604). <https://doi.org/10.1145/3375462.3375528>
- Kember, D., McKay, J., Sinclair, K., & Wong, F. K. Y. (2008). A four-category scheme for coding and assessing the level of reflection in written work. *Assessment & Evaluation in Higher Education*, 33(4), 369–379. <https://doi.org/10.1080/02602930701293355>
- Kovanović, V., Joksimović, S., Mirriahi, N., Blaine, E., Gašević, D., Siemens, G., & Dawson, S. (2018). Understand students' self-reflections through learning analytics. In *Proceedings of the 8th international conference on learning analytics and knowledge* (pp. 389–398). <https://doi.org/10.1145/3170358.3170374>.
- Krippendorff, K. (2003). *Content analysis: An introduction to its methodology* (p. 8). Sage.
- Kyllonen, P. C. (2012). Measurement of 21st century skills within the common core state standards. In *Invitational research symposium on technology enhanced assessments* (p. 24).
- Lai, E. R., & Viering, M. (2012). *Assessing 21st century skills: Integrating research findings*. Pearson.
- Lee, Y.-W., & Sawaki, Y. (2009). Cognitive diagnosis approaches to language assessment: An overview. *Language Assessment Quarterly*, 6(3), 172–189. <https://doi.org/10.1080/15434300902985108>
- Martin, C. K., Nacu, D., & Pinkard, N. (2016). Revealing opportunities for 21st century learning: An approach to interpreting user trace log data. *Journal of Learning Analytics*, 3(2), 37–87.
- Milligan, S. (2020). Standards for developing assessments of learning using process data. In M. Bearman, P. Dawson, R. Ajjawi, J. Tai, & D. Boud (Eds.), *Re-imagining university assessment in a digital world* (Vol. 7, pp. 179–192). Springer. https://doi.org/10.1007/978-3-030-41956-1_13
- Milligan, S. K., & Griffin, P. (2016). Understanding learning and learning design in MOOCs: A measurement-based interpretation. *Journal of Learning Analytics*, 3(2), 88–115. <https://doi.org/10.18608/jla.2016.32.5>
- Mislevy, R. J., Behrens, J. T., Dicerbo, K. E., & Levy, R. (2012). Design and discovery in educational assessment: Evidence-centered design, psychometrics, and educational data mining. *Journal of Educational Data Mining*, 4(1), 11–48.
- Nederhof, A. J. (1985). Methods of coping with social desirability bias: A review. *European Journal of Social Psychology*, 15(3), 263–280. <https://doi.org/10.1002/ejsp.2420150303>
- Pellegrino, J. W. (2017). *Teaching, learning and assessing 21st century skills* (pp. 223–251). <https://doi.org/10.1787/9789264270695-12-en>
- Pongpaichet, S., Nirunwiroj, K., & Tuarob, S. (2022). Automatic assessment and identification of leadership in college students. *IEEE Access*, 1. <https://doi.org/10.1109/ACCESS.2022.3193935>
- Quinn, R. E. (1988). *Beyond rational management: Mastering the paradoxes and competing demands of high performance* (pp. xxii, 199). Jossey-Bass.
- Rios, J. A., Ling, G., Pugh, R., Becker, D., & Bacall, A. (2020). Identifying critical 21st-century skills for workplace success: A content analysis of job advertisements. *Educational Researcher*, 49(2), 80–89. <https://doi.org/10.3102/0013189X19890600>

- Rohs, F. R., & Langone, C. A. (1997). Increased accuracy in measuring leadership impacts. *Journal of Leadership Studies*, 4(1), 150–158. <https://doi.org/10.1177/107179199700400113>
- Rupp, A. A., & Templin, J. L. (2008). Unique characteristics of diagnostic classification models: A comprehensive review of the current state-of-the-art. *Measurement: Interdisciplinary Research and Perspectives*, 6(4), 219–262. <https://doi.org/10.1080/15366360802490866>
- Rupp, A. A., Templin, J., & Henson, R. A. (2010). *Diagnostic measurement: Theory, methods, and applications*. Guilford Press.
- Salleb-Aouissi, A., Vrain, C., & Nortel, C. (2007). *QuantMiner: A genetic algorithm for mining quantitative association rules* (pp. 1035–1040).
- Sartori, L., & Theodorou, A. (2022). A sociotechnical perspective for the future of AI: Narratives, inequalities, and human control. *Ethics and Information Technology*, 24(1), 4. <https://doi.org/10.1007/s10676-022-09624-3>
- Shum, S. B., & Crick, R. D. (2016). Learning analytics for 21st century competencies. *Journal of Learning Analytics*, 3(2), 6–21. <https://doi.org/10.18608/jla.2016.32.2>
- Sondergeld, T. A., & Johnson, C. C. (2019). Development and validation of a 21st century skills assessment: Using an iterative multimethod approach. *School Science and Mathematics*, 119(6), 312–326. <https://doi.org/10.1111/ssm.12355>
- Sutherland, M. A., Amar, A. F., & Laughon, K. (2013). Who sends the email? Using electronic surveys in violence research. *The Western Journal of Emergency Medicine*, 14(4), 363–369. <https://doi.org/10.5811/westjem.2013.2.15676>
- Ullmann, T. D. (2019). Automated analysis of reflection in writing: Validating machine learning approaches. *International Journal of Artificial Intelligence in Education*, 29(2), 217–257. <https://doi.org/10.1007/s40593-019-00174-2>
- Vockley, M. (2007). Maximizing the impact: The pivotal role of technology in a 21st century education system. In *Partnership for 21st century skills*. Partnership for 21st Century Skills. <https://eric.ed.gov/?id=ED519463>
- Wu, Y., & Crocco, O. (2019). Critical reflection in leadership development. *Industrial and Commercial Training*, 51(7/8), 409–420. <https://doi.org/10.1108/ICT-03-2019-0022>