







An Interactive Recommender System Based on Reinforcement Learning for Improving Emotional Competences in Educational Groups

Eleni Fotopoulou^{1,2}(✉) , Anastasios Zafeiropoulos¹ , Michalis Feidakis² ,
Dimitrios Metafas², and Symeon Papavassiliou¹ 

¹ Institute of Communication and Computer Systems,
National Technical University of Athens, 15773 Zografou, Greece
efotopoulou@netmode.ntua.gr

² University of West Attica, 12241 Egaleo, Greece

Abstract. The development of Social and Emotional competences of students can significantly improve their learning and social outcomes. This prompts for tools to assist tutors in accomplishing social and emotional learning activities and evaluating the impact achieved. To do so, the blending of Recommender Systems with Machine Learning technologies can be proven beneficial for the design of intelligent and self-learning tools with the capacity to recommend activities, aligned with the social and emotional needs of educational groups. In the current manuscript, we detail a modeling approach for an interactive Recommender System that aims to suggest educational activities to tutors for improving the social and emotional competences of students, taking advantage of Reinforcement Learning techniques. A Reinforcement Learning model has been designed that considers the evolution of students' social and emotional characteristics and the provided feedback through a set of interactions. Short evaluation of the detailed approach is provided, focusing on validating its appropriateness to serve educational needs.

Keywords: Reinforcement Learning · Social and emotional competences · Interactive Recommender System · Education

1 Introduction

The improvement of Social and Emotional Competences of both students and tutors is associated with positive learning and social interaction outcomes, such as self-efficacy, problem solving, positive classroom climate, stress management, conflict resolution and prevention of youth behavioral problems [3, 13, 19]. Despite the underestimate inclusion of respective educational activities in the past, Social-Emotional Learning (SEL) have drawn attention during the last years deploying structured emotional education activities in existing curricula [20].

Nevertheless, the adoption of such activities is not straightforward due to the heterogeneity of methodologies for conducting SEL activities, the existence of scattered and not always evaluated material and the inadequate -in some cases- expertise on behalf of the tutors [21]. Thus, there is a need to develop tools and methodologies to facilitate tutors to accomplish educational interventions targeted to social and emotional characteristics of students, as well as evaluate their impact at individual and group level.

Towards this direction, the exploitation of emerging technological solutions in the areas of Recommender Systems (RS) and Machine Learning (ML) seems promising. The blending of these technologies can enable the provision of targeted recommendations, with respect to the group's social and emotional characteristics and identified needs, as well as continuous feedback by students and tutors. By taking advantage of novel ML techniques, such as Reinforcement Learning (RL), the outcomes of the provided recommendations and the dynamics of interaction among students can be modeled, monitored and periodically evaluated, leading to improvements in the efficiency of the RS.

In the current manuscript, we detail an environment for an interactive RS, able to provide recommendations to tutors, taking advantage of the implementation of a RL model. The latter incorporates concepts related to (i) the modeling of the educational group's social and emotional profile (e.g., emotional awareness, conflict resolution, self-efficacy), (ii) the type and the characteristics of the activities (e.g., targeted to competences such as emotional consciousness, emotional regulation), (iii) the selection of the activity per step (e.g., matching of activities with the identified needs), and (iv) the type of the response per recommendation (e.g., acceptance ratio, activity quality, positive interactions among students). The respective environment has been built upon RecSim [7] that regards a release of a configurable platform for authoring simulation environments for RS that naturally supports sequential interaction with users. Recsim is used as a basis for the models' definition and the generation of training data. Short evaluation results and directions for future research are provided.

2 State of the Art Analysis and Motivation

2.1 Social and Emotional Competences Evaluation

The first step towards the development of mechanisms for strengthening social and emotional competences within educational groups deals with the homogeneous representation of the collected information based on the adoption or adaptation of relevant Emotional Intelligence (EI) models. By EI, we refer to the capability of individuals to recognize their own emotions and those of others, discern between different feelings and label them appropriately, use emotional information to guide thinking and behavior, manage and adjust emotions to adapt to environments or achieve one's goal(s). Three main EI models are considered, namely the Ability [15], the Mixed [6], and the Trait model [17].

In all cases, definition of EI and interlinking of EI indicators with specific emotional competences -taught or improved through training activities- is not

strict or standardized. An interesting work for expressing emotional competences as a set of micro-competencies that can be assessed and evaluated is based on the model provided in [2]. In this model, emotional competences are expressed in five main dimensions, namely emotional consciousness, emotional regulation, emotional autonomy, social competences and competences for life and well-being.

Sociometry oriented models are also considered [5, 16], where focus is given on qualitative methods for measuring metrics related to peer relationships. A sociogram can be drawn on the basis of many different criteria such as peer relations, channels of influence, lines of communication and can be used to identify pathways for social acceptance for misbehaving members.

2.2 Interactive Recommender Systems in Education

Traditional RS are broadly categorized in three areas; content-based, collaborative filtering and hybrid systems [10, 12]. Content-based recommendations are based on individual user characteristics and preferences, considering user's former selections. Collaborative filtering recognizes commonalities among users or items and recommends preferred items by similar users. Hybrid RS combine the aforementioned two approaches, aiming at the selection of the best algorithms for achieving greater efficiency.

All three categories have been applied in the field of education, especially for suggesting effective e-learning paths to students, supporting teaching activities and enhancing academic performance [12, 18]. Based on a mapping study to investigate the use of RS in education [18], the most commonly applied approach is the hybrid one, followed by the collaborative filtering approach. Open research areas include the introduction of artificial intelligence in RS algorithms to improve personalization of academic choices and the need for consideration of differences in the learner profile and characteristics [18]. Similar limitations are identified in the aforementioned types of RS, independently of their application domain. Traditional RS provide a relatively static list of recommendations, while the collected feedback is usually limited to the acceptance or not of the recommendation [1]. Also, their main focus is on estimating immediate user engagement without considering the long-term effects on user behavior [9].

A recent trend in RS is Interactive RS, where the user is able to interact with the provided recommendation and give feedback that may impact the results in real-time [1]. Appropriate modeling of the dynamics of user interaction is crucial for developing mechanisms able to improve users' long-term engagement and overall satisfaction, upon a sequence of interactions [7]. The latter entails exploitation of advances in ML mechanisms and mainly in RL [7, 9]. In this way, richer forms of mixed-mode interactions, supporting a variety of system actions (e.g., provision of endorsements to students, preference elicitation) and user responses (e.g. indirect/direct feedback) can be applied.

2.3 Motivation

The design of our approach is dictated by the need to exploit capabilities of novel RS towards the improvement of social and emotional competences of users in real educational groups. The blending of RS with RL is considered promising for the development of solutions able to provide effective recommendations by considering the continuous mixed-mode interactions and the feedback provided on behalf of the users. Our aim is to provide a theoretical model along with a proof-of-concept implementation -in a simulation environment in Recsim [7]- that can be adopted and appropriately adapted to fit various educational settings.

3 Interactive Recommender System Design

In the proposed approach (Fig. 1), a RS interacts with a group of users within a classroom and provides recommendations for the implementation of educational activities, aiming at the improvement of the group’s social and emotional competences. Decision making in the interactive RS is supported by a RL model.

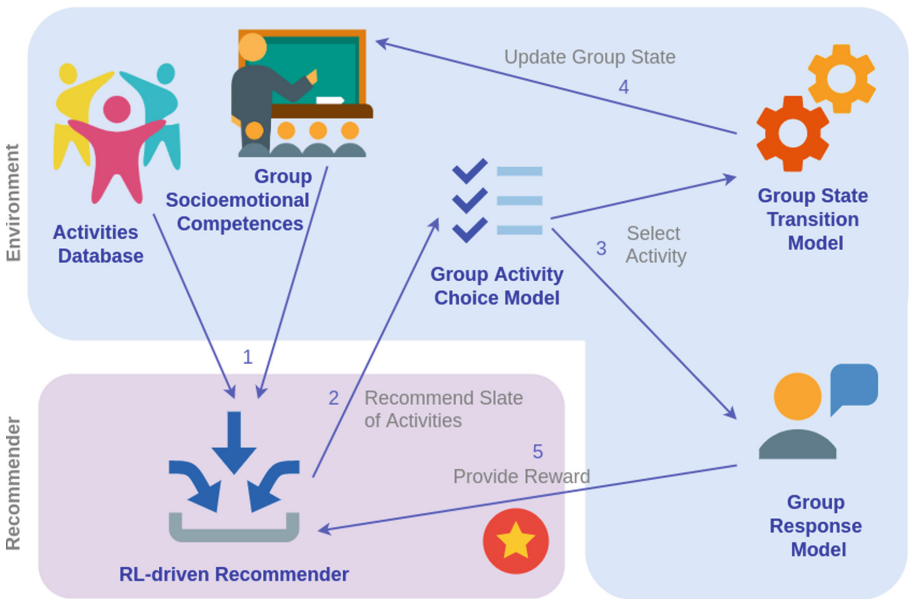


Fig. 1. Interactive recommender system high level view.

The main specified concepts of the RL model are (i) the Activities, (ii) the Group Social and Emotional Competences, (iii) the Group Activity Choice Model, (iv) the Group Response Model, and (v) the Group State Transition Model. The Activities include social and emotional education activities, made

available to the RS through a database. Each Activity is related to a set of competences that can be improved within the educational group. These competences are represented on the Group Competences Model. Based on the provided set of Activities and the snapshot of the Group Competences (step 1), a subset of activities -known as slate- is recommended (step 2). From this slate, a specific activity is selected, based on the business logic supported by the Group Activity Choice model (step 3). Next, the selected activity is accomplished, under the supervision of the tutor of the group. Continuous feedback on behalf of the tutor and the students is collected with regards to the acceptance, the applicability and the attractiveness of the implemented activity, leading to the formulation of the Group Response, while in parallel the educational group state may be changed, based on a defined Group State Transition model (step 4). The latter evaluates the transition in the group's social and emotional state, considering the learning impact of the applied activity and the evolution of the group's emotional competences, within some period of time. Based on the achieved effectiveness towards the goals set, a reward is provided (step 5) as a feedback from the environment, which can be consumed on behalf of any RL agent.

3.1 Social and Emotional Activities

Each activity is classified in accordance with a list of social and emotional competences it can improve. The exact type of the addressed competencies is associated with the adopted social and emotional model, considering the models detailed in Sect. 2.1. In the current work, we have used the five dimensions of the emotional competences model defined in [2]. These dimensions are decomposed at thirty fine-grained micro-competences, used for characterizing each of the activities (Fig. 2). Each activity may address more than one micro-competence at the same time. A value ranging from $[-1..1]$ is assigned to each micro-competence, denoting whether it is within the main learning target (value close to 1) or outside the learning scope (value close to -1) of the activity. It should be noted that the provided environment is modular, supporting the selection of a different EI or Emotional Competences model with small adaptation effort.

In addition to the characterization of activities in terms of targeted social and emotional features, further characteristics, such as the quality and the duration of each activity, are considered. The quality indicates the effectiveness of the activity towards improving the targeted micro-competencies and is important input for the Group Response model. The duration is also crucial since a set of activities have to be accommodated within a specific time budget. Upon the implementation of each activity, the overall time budget is reduced by the associated duration. All the activities are made available through a database. Based on the available pool of activities, part of them are selected at each step, formulating the suggested activities slate.

For ease of comprehension, a sample educational activity -tackling the micro-competences of respect and collaboration as shown in Fig. 2- includes a theoretical introduction on the importance of these competences and a consequent split

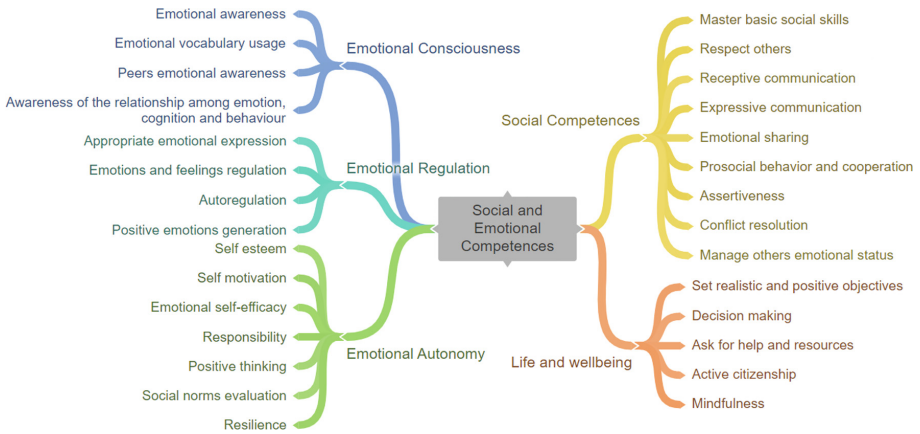


Fig. 2. Social and emotional competences classification [2].

of the group on sub-groups, where the students are asked to create a choreography for a specific song in a certain amount of time. The students have to present their outcome to the rest of the team and reflect on their behaviour at the end of the session.

3.2 Educational Group Social and Emotional State

The educational group social and emotional state is represented based on the same model applied for classification of the activities. Personal and group metrics are estimated, considering individuals’ social and emotional state, as well as the interaction among the group members. A value ranging from $[-1..1]$ is assigned to each micro-competence, indicating whether it needs to be strengthened (value close to 1) or it is well developed (value close to -1).

In the current work, the group’s state is used for the provision of recommendations and the evaluation of their efficiency. The objective is to offer activities that match the identified social and emotional needs of the group. High values of a micro-competence in the group profile reflect the need to improve it by conducting high-quality activities targeting this micro-competence.

In addition to the social and emotional needs of the educational group, further characteristics are specified and monitored. These include the group’s time budget (e.g., 30 lecture hours), the set of previously applied and preferred activities by the educational group, the probability that an activity is appealing to the tutor, the educational group receptiveness level upon the completion of an activity and the step penalty applied in case no activity is selected.

3.3 Group Activity Choice, Response and State Transition Model

The Group Activity Choice model supports the selection of the activity per recommended slate that -in a real educational environment- is done by the tutor

of the educational group. The rejection of all the proposed activities in a specific slate is also a valid option but includes a penalty at the group's time budget. Two choice models are made available, namely the multinomial logit [14] and the exponential cascade model [11], as they are provided by Recsim [7].

In the multinomial logit model, multinomial logistic regression is used to predict the choice probabilities for all the offered activities of the slate. A group g selects an activity a , with an unnormalized probability $I(g, a)$ that is given by the dot product $I(g, a) = \mathbf{g}\mathbf{a}$. This value corresponds to the capacity of the activity to tackle the group's social and emotional needs and is called as *group's interest* for the given activity. The normalized probability of group g selecting an activity a from a slate of activities A is given in Eq. 1.

$$P(a|A) = \frac{e^{I(g,a)}}{\sum_{j \in A} e^{I(g,j)}} \quad (1)$$

In the exponential cascade model, it is supposed that, in each iteration, the tutor pays attention to one activity of the slate at a time, with exponential decreasing focus as the tutor moves down the slate. The probability that an activity in position i of the slate is checked by the tutor is modeled as $\beta_0\beta^i$, where β_0 is a base probability and β the decay value from activity i to $i + 1$. Both of them are constant values in the range of $[0, 1]$. The probability that the tutor of the group g selects an activity a in the position j within the slate A of activities is given in Eq. 2.

$$P(a|A) = \beta_0\beta^j I(g, a) \quad (2)$$

The Group Response model represents the response provided on behalf of the group, upon the completion of an activity. The response includes information regarding the positive interactions among the group members, the perceived quality of the activity on behalf of the group, the activity duration, the evaluation of the tutor as activity facilitator by the group members and the self-evaluation of the tutor. The perceived quality of the activity on behalf of the group composes the reward that the environment returns to the RS per execution step. In the current work, the reward is modeled as the multiplication of the group's interest $I(g, a)$ for the realized activity a with the activity quality a_q and the group receptiveness level g_r , as presented in Eq. 3.

$$R_a = I(g, a)a_qg_r \quad (3)$$

The Group State Transition model evaluates the new state of the social and emotional profile of the educational group upon the realization of an activity. The transition probability $P(s_{i+1}|s_i, A)$ reflects the probability that the group features are updated from state s_i to state s_{i+1} upon the completion of activity A . The main task of the Group State Transition model is to slightly update the group's socio-emotional competences. Depending on the group's interest for the selected activity, the values of the associated competences are increased or decreased by a step. When the group's interest is more than a threshold, it is

supposed that the group’s emotional needs for training at the specific micro-competences is addressed satisfactorily and the group’s values are decreased for the specific micro-competences. In a similar way, when the group’s interest is below a specific threshold, the group’s state values are increased, denoting a need for addressing these competences in upcoming iterations. The Group State Transition model can be also “rigid” so that the group state never changes during the episode execution.

4 Implementation and Initial Evaluation Results

4.1 Educational Activities and Group Data Preparation

The input data for the evaluation of the proposed approach is provided based on sampling techniques for both the features of the education groups and the activities. Attention has been given on supporting various combinations in terms of the competencies tackled by the activities and the competences that need strengthening within the group. For the group features, we provide uniformly distributed samples from $g \in [-1, 1]^F$, where F is the number of micro-competences. Similarly, for the activities features, we provide uniformly distributed samples from $a \in [-1, 1]^F$. For the quality of the activities, two ways of sampling are applied. At the Basic Activity sampling, a static quality of $q = 1.0$ is given to all activities, aiming to leave the group’s social and emotional characteristics to fully drive the selection of an activity. At the Utility Based Activity sampling, 70% of the activities are considered of low quality and follow a linear distribution from $[-3, 0]$, while the rest 30% is considered of high quality and follows a linear distribution from $[0, 3]$. This is a more challenging setup for the RS, since it has to select activities of good quality that cover specific micro-competences.

4.2 Evaluation Results

To realize an evaluation over indicative educational settings, an environment has been developed within Recsim and made openly available at [4]. In this section, some preliminary evaluation results are provided, based on the instantiation of the modeled environment according to the data sampling detailed in Subsect. 4.1 and the group models and reward function detailed in Subsect. 3.3. The main objective is to validate the expressive and accurate modeling of the provided approach and its capacity to onboard multiple experimentation scenarios in the future.

The recently published Q-value algorithm, known as SLATEQ [8] has been used as an RL algorithm. SLATEQ renders RL tractable with slates. It is a new slate decomposition technique that estimates the long-term value (LTV) of a slate of items by directly using the estimated LTV of the individual items on the slate. The simulation environment has been configured with 100 groups, a pool of 7 activities per step and 3 activities for each slate. The training period

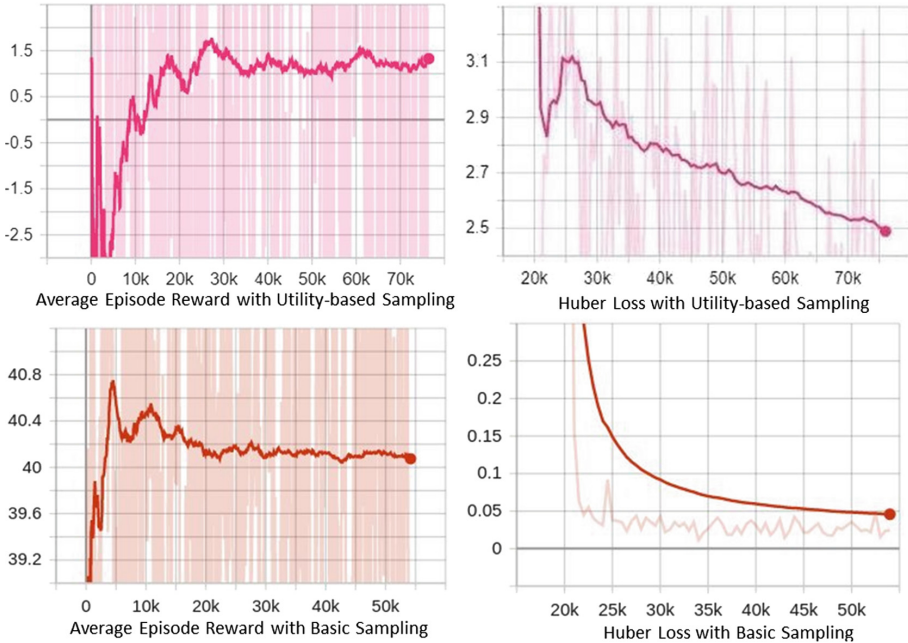


Fig. 3. Average reward and Huber loss per environment setup.

for each educational group is set to 20 lecture hours, while the penalty step to 15 min. 30 social and emotional microcompetences are considered. The number of training steps is about 60K with 30 maximum evaluation steps per episode.

The average episode reward and the Huber loss in case of two environment setups are depicted in Fig. 3. Utility based sampling has been used in the first case (first row in the figure), opposed to basic activity sampling (with a stable quality for all documents) in the second case. In both cases, the average reward is following an increasing trend, validating that the RL agent is learning from the feedback received by the environment. Larger improvement in the learning process -in absolute terms- is accomplished in the first case. In both cases, the RS was able to reduce the Huber losses during the training execution.

5 Conclusions and Future Research Areas

In the current manuscript, we have detailed an approach for a RL-driven interactive RS that provides recommendations for educational activities, aiming at the improvement of social and emotional competences of the group members. The modeling of the approach is considered indicative for developing solutions that provide interactive recommendations to students in various settings. Initial evaluation results are presented for the validation of the proposed approach.

A set of open research areas have been identified, including the design of mechanisms for decomposition of the educational group state into the students'

personal state, the extensive evaluation in various settings and the development of RL agents that provide optimal reward convergence within less training steps.

Acknowledgements. This research effort is supported by ICCS Research Award under Grant Number 65020603.

References

1. Alkan, O., Daly, E.M., Botea, A.: An evaluation framework for interactive recommender systems. In: UMAP 2019 Adjunct, New York, NY, USA, pp. 217–218 (2019)
2. Bisquerra Alzina, R., Pérez González, J.C., García Navarro, E.: *Inteligencia emocional en educación*. Editorial Síntesis (2015)
3. Feidakis, M.: A review of emotion-aware systems for e-learning in virtual environments. In: Caballé, S., Clarisó, R. (eds.) *Formative Assessment, Learning Data Analytics and Gamification*, pp. 217–242. Academic Press, Boston (2016)
4. Fotopoulou, E., Zafeiropoulos, A.: RL-driven interactive recommender based on socioemotional behavioural models (2020). <https://gitlab.com/netmode/arrle>
5. Fotopoulou, E., Zafeiropoulos, A., Alegre, A.: Improving social cohesion in educational environments based on a sociometric-oriented emotional intervention approach. *Educ. Sci.* **9**(1) (2019). <https://doi.org/10.3390/educsci9010024>
6. Goleman, D.: *Emotional intelligence: why it can matter more than IQ*. Bantam (1996)
7. Ie, E., et al.: RecSim: a configurable simulation platform for recommender systems (2019)
8. Ie, E., et al.: SlateQ: a tractable decomposition for reinforcement learning with recommendation sets. In: *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, China*, pp. 2592–2599 (2019)
9. Ie, E., et al.: Reinforcement learning for slate-based recommender systems: a tractable decomposition and practical methodology (2019)
10. Isinkaye, F., Folaajimi, Y., Ojokoh, B.: Recommendation systems: principles, methods and evaluation. *Egypt. Inf. J.* **16**(3), 261–273 (2015)
11. Joachims, T.: Optimizing search engines using clickthrough data. In: *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, NY, USA*, pp. 133–142 (2002)
12. Khanal, S., Prasad, P., Alsadoon, A., Maag, A.: A systematic review: machine learning based recommendation systems for e-learning. *Educ. Inf. Technol.* 1–30 (2019). <https://doi.org/10.1007/s10639-019-10063-9>
13. Linares, L., et al.: Developing cognitive-social-emotional competencies to enhance academic learning. *Psychol. Schools* **4**, 405–417 (2005)
14. Louviere, J.J., Hensher, D.A., Swait, J.D., Adamowicz, W.: *Stated Choice Methods: Analysis and Applications*. Cambridge University Press, Cambridge (2000)
15. Mayer, J.D., Caruso, D.R., Salovey, P.: The ability model of emotional intelligence: principles and updates. *Emot. Rev.* **8**(4), 290–300 (2016)
16. Moreno, J.: *Who shall survive? Foundations of sociometry, group psychotherapy and socio-drama*, January 1953
17. Petrides, K.: Ability and Trait Emotional Intelligence. *The Blackwell-Wiley Handbook of Individual Differences*, pp. 656–678, March 2013

18. Rivera, A.C., Tapia-Leon, M., Lujan-Mora, S.: Recommendation systems in education: a systematic mapping study. In: Rocha, Á., Guarda, T. (eds.) ICITS 2018. AISC, vol. 721, pp. 937–947. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-73450-7_89
19. Salovey, P., Sluyter, D.J.: Emotional Development and Emotional Intelligence: Educational Implications. Basic Books (1997)
20. Schonert-Reichl, K.A.: Advancements in the landscape of social and emotional learning and emerging topics on the horizon. *Educ. Psychol.* **54**(3), 222–232 (2019). <https://doi.org/10.1080/00461520.2019.1633925>
21. Sokal, L., Katz, J.: Social emotional learning and inclusive education. The Oxford Encyclopedia of Education, January 2017