# Towards Adaptive Learning Systems: The Potential of Reinforcement Learning in eLearning systems

Frederik Willemsen

May 21, 2026

# Contents

# 1. Introduction

With the uprising of improving Artificial Intelligence algorithms and more available datasets to train these algorithms on, the implementation of Reinforcement learning into eLearning systems has become a recent trend. E-learning systems gained enormous popularity during the COVID-19 pandemic when schools had to be closed, and people had to go online in order to educate themselves. Duolingo and Babbel are examples of e-learning systems used to learn languages, but many more platforms, like Moodle or Kahoot, exist where you can start to educate yourself online. As platforms like Duolingo have millions of users, the question arose of how to optimize the learning experience for each of the users, to make learning more fun, easy, and accessible. Reinforcement Learning offers a lot of possibilities to tackle these problems, with advanced algorithms and a lot of data that is collected on these online platforms such as in Duolingo [Bicknell et al.(2023)Bicknell, Brust, and Settles], it might just be a perfect solution to make learning more fun and personalized. Traditionally, learning and especially education have been a system that was stuck without progress. The same answers and the same questions are provided to people with different levels of knowledge, may this be the most optimal way to learn? Unlikely! This article is a review of the potential Reinforcement Learning has in eLearning Systems, to make learning more adaptive, efficient, and fun!

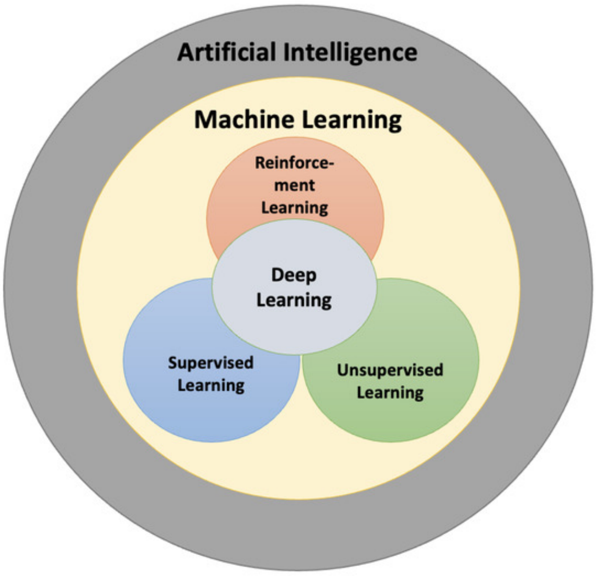


**Figure 1** Overview of Artificial Intelligence and its subfields, taken from [Mon et al.(2023)Mon, Wasfi, Hayajneh, Slim, and Abu Ali].

# 2. Background

## 2.1. eLearning and adaptive Systems

E-Learning Systems are defined as all forms of online electronic-supported systems that enable the opportunity to learn and teach. Furthermore, e-learning Systems should aid the construction of knowledge, practice, and individual experience of the learner. Thereby, the electronic devices used serve as media to implement the learning process. What makes an e-learning system adaptive is its capability to track the learner's knowledge, style, preferences, and behaviors over time. With this learned data, an adaptive system would be able to use AI/ML to tailor the content to the Learner and suggest an optimal learning path, analyze weaknesses, adjust the learning process and difficulty to the Learner's level in real time [Sweta(2021)].

## 2.2. Basics of Reinforcement Learning

Reinforcement Learning is a method in machine learning. It focuses on how agents can learn to make decisions through interaction with an environment to maximize the cumulative reward function over time. While Supervised learning relies on labeled data, and Unsupervised Learning uncovers hidden patterns, Reinforcement learning learns optimal actions through trial and error without explicit instructions **??**. Reinforcement Learning therefore suffers under the exploration-exploitation trade-off, which refers to the decision dilemma between exploring new states which could be less optimal than a known action and exploiting an already known action that is guaranteed to return the best reward under all the currently known actions [Sutton and Barto(2018)]. Reinforcement Learning involves multiple core elements. The State is the part of the environment at a particular moment. This represents the situation the agent is in at the moment. At any state, the agent can decide on multiple Actions that he can take. For any action taken in a state, the agent receives a Reward, which tells the agent how good or bad his action was. The decision process of which action to take is also influenced by the rewards of future actions. However, these are controlled by the discount factor. Intuitively, the discount factor controls how important future rewards are compared to immediate rewards [Ghasemi and Ebrahimi(2024)].

## 2.3. Pedagogical Theories

Adaptive eLearning systems are the most effective when they support established theories from pedagogical research and are focused on the learner and their user experience. In Constructivism, learners should build knowledge through experience and interaction, as learning is seen as an active process; this aligns with adaptive technologies such as eLearning systems [Tam(2000)]. Additionally, Mastery Learning states that given enough time and effort, all learners can achieve a high level of understanding [Guskey(2007)]. Reinforcement Learning can help here by adjusting the difficulty.

# 3. Reinforcement Learning in Education

## 3.1. Use cases of RL Applications in eLearning

Reinforcement learning can be applied across a wide range of tasks in e-learning. A few of the potential tasks where Reinforcement Learning is studied to be used are presented here: Curriculum Sequencing is one of the most popular use cases of RL algorithms in eLearning. The algorithm is used to determine the most effective sequence in which the learning material should be presented to the student to maximize learning gains and make learning most effective.
The paper [Rafferty et al.(2011)Rafferty, Brunskill, Griffiths, and Shafto] explored how to optimize teaching with Reinforcement Learning, by modeling it as a POMDP.

Another use case is strategies where adaptive feedback is provided to the student. Reinforcement Learning is used to adapt feedback timing and content based on previous behaviour and interaction of the student. The goal is to help learners with the right information at the right time. This was studied by [Chi et al.(2011)Chi, Vanlehn, Litman, and Jordan], the authors explained how Reinforcement Learning driven feedback improves problem solving in physics tutoring systems.

On top of that, dynamic difficulty adjustment can be done by the RL algorithm. The goal is to adapt the difficulty level according to the progress the student is making in real-time to match the student's ability level, therefore reducing frustration and enhancing motivation and engagement.
In [Rahimi et al.(2023)Rahimi, Moradi, Vahabie, and Kebriaei], RL-tuned visual memory games led to better engagement and outcomes.

Spaced Repetition is also an interesting use case where reinforcement learning can help. The trained Reinforcement Learning agents are optimized to suggest when to review learning material for the user, such that long-term memorization is maximized.

## 3.2. Techniques used

A multitude of different reinforcement learning strategies are used in e-learning, which depend on the context and the intended implementation, as well as various design choices. In most systems, Reinforcement Learning is implemented as a Markov Decision Process, which models learners' progress as a sequential process from state to state. Techniques such as Q-Learning and Deep Q-Networks are most effective for modeling the states. At the same time, multi-armed bandit algorithms are used when immediate feedback and short-term planning, such as choosing optimal reminders or learning tips, is the goal. On top of that, offline RL methods are used for educational purposes, as safety and ethical constraints 5.2 are important.

## 3.3. Rule-based vs. Data-driven Systems

Legacy eLearning systems rely on rule-based systems, which are designed in an if-then-else manner. This effectively treats all learners similarly and assumes everyone is starting from the same knowledge level, and does not provide personalized support features. Data-driven methods use machine learning or reinforcement learning to tailor feedback for the user; however, classical machine learning models are dependent on the goodness of their training data. Reinforcement Learning is therefore used to improve these systems while deployed on real interaction data. They dynamically learn optimal decisions, based on long-term learning outcomes, and then suggest the best decision for each learner. Research has shown that data-driven systems often outperform rule-based systems. A study about adaptive Learning in higher education [du Plooy et al.(2024)du Plooy, Casteleijn, and Franzsen] showed that personalized feedback increased academic performance and student engagement.

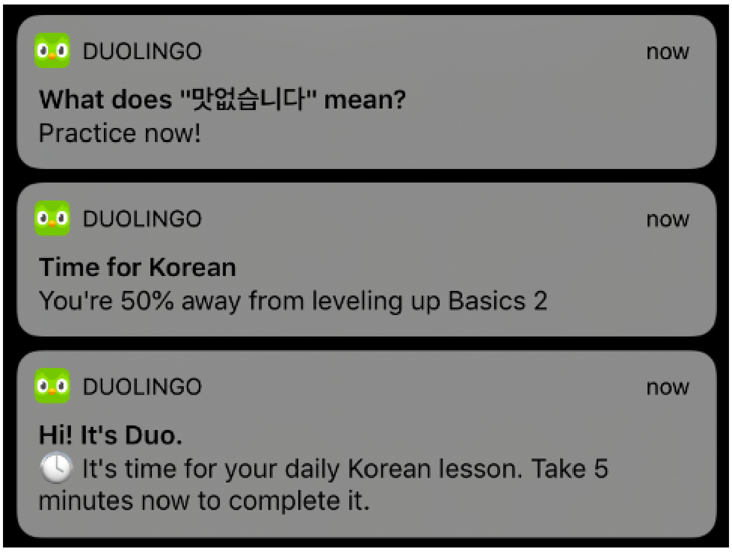


**Figure 2** Example for push notifications received by Duolingo, taken from [Yancey and Settles(2020)].

| **Bandit Algorithm** | **Avg. Reward ($\bar{r}$)** | **Rel. Diff.** |
|---|---|---|
| | $\pm 0.00015$ | |
| Baseline (random) | 0.1295 | - |
| Template | 0.1311 | +1.2% |
| Template+UI Language | 0.1318 | +1.8% |

**Table 1** Variations of the bandit algorithm and compared to a baseline, for push notifications, taken from [Yancey and Settles(2020)].

# 4. Existing Research

## 4.1. Case Studies

In the paper [Yancey and Settles(2020)], the author describes how a multi-armed bandit algorithm, which is a type of RL algorithm, is used in Duolingo to receive personalized push notifications, as can be seen in 2. These daily push notifications are selected by multi-armed bandit systems to maximize engagement with the App. The algorithm is context-aware, which means that a User with a daily streak would therefore be more likely to get a notification to keep their streak than a user who does not have a streak. Furthermore, the algorithm introduced a recency penalty, meaning that the reminders are getting more effective the longer they have not been used. The deployment of these measures led to a measurable increase in user engagement, lesson completions, and daily active users. The Reinforcement Learning algorithm used there works in production and is used daily to suggest automated messages for millions of Duolingo users. Testing of these hypotheses showed that 0.4% more lessons were completed, and the new-user retention increased by approximately 2%, compared to the baseline, as can be seen in 1.

In another paper, [Azhar et al.(2022)Azhar, Segal, and Gal], another promising idea is presented. The author implemented a Reinforcement Learning pipeline, which personalizes the order of questions for each learner. He achieved this by framing the problem as a Markov Decision Process, where states represent the student's knowledge, actions are the selection of the next questions, and the rewards are based on correctness and engagement of the learner. For the training and testing process, he used the large-scale student interaction data (EdNet dataset), without live deployment. The pipeline, therefore, was not deployed to a live simulation for students, but adapted and trained offline. This paper serves as a proof-of-concept, showing how such a pipeline could work and which design principles matter for performance and generalisation. With his research the author demonstrated that Reinforcement Learning policy can be used effectively in eLearning systems, under the assumption that state representations are well-designed. The paper also showed the importance of feature engineering in educational Reinforcement Learning systems.

On top of that the paper [Fahid and Chi(2021)] looks at how applied deep reinforcement learning can be used to provide feedback, hints, or explanations in an intelligent tutoring system. The system seeks to learn when students need help during their assignments, and then gives feedback at the right moment to maximize cognitive learning. Furthermore, the system seeks to creatively explain ideas, not just clicking and guessing, which should enhance user engagement and deeper thinking about the subject. It learns from past interactions to decide on the best moment when to offer guidance, instead of following strict intervals. The system was also trained on real student data, making the data real and based on real human behaviour. The results of the study showed that the system did a better job at providing feedback and supporting students compared to a human tutor who follows preset strategies. This showcased how smart Machine learning methods can outperform human actions, by adaptively providing help only when needed.

# 5. Challenges and Limitations

The large-scale deployment of Reinforcement Learning has already begun with leading companies like Duolingo implementing it in their algorithms to optimize push notifications. In more traditional domains in education it is still more difficult to implement it.

## 5.1. Deployment Challenges

RL algorithms return better results on larger data samples than on smaller ones. This is a result of the exploration-exploitation trade-off; if there are fewer states to explore, the algorithm will find the best option available faster, but the overall results might still be suboptimal. RL algorithms, like all data-driven methods, are extremely data hungry; therefore, it is also extremely important how the states are modeled for optimal results. In educational institutions, harnessing a huge amount of data might clash with privacy concerns 5.2. Effective data collection might also be a problem in developing countries, which can lead to worse results [Dake(2023)].

## 5.2. Privacy and Ethical Concerns

Machine Learning methods are hungry for data; they need it to improve the models and be more accurate. If you interact with a chatbot, your data will be saved and processed to improve the algorithms. Not only that, but your data will also be sold to third parties; you actually do not own your own data on some online platforms, where you actually transfer your copyright when uploading something. Digital assistants like Amazon's "Alexa" have been shown to be very exploitable to privacy risks and listen even when deactivated [Bartneck et al.(2021)Bartneck, Lütge, Wagner, and Welsh]. In the context of e-learning, this also raises huge issues. While interacting with online Learning solutions, the algorithms learns your habits, it also learns how you respond to certain actions and your consumer behaviour. Your actions will be internalized in the system and used to improve the results for other people. Ultimately, this can lead to a surveillance that should not be expected when interacting with an e-learning system. If these systems collect enough data about one, this potentially leads to identity theft or fraud. The article [Paul(2024)] outlines several security concerns and common privacy breaches by AI. Oftentimes, these include data breaches, such as weak security protocols or third-party vendors, but also misuse and inadequate control of data. The misuse can lead to unauthorized access by computer algorithms to protected information. Other problems include the bias and inference risk these systems have on their training data, which can lead to worse results for minorities and favor certain demographics over others [Yulianti(2025)].

# 6. Opportunities and Future Directions

## 6.1. Hybrid Models

In [Ma(2023)], the author presented the idea of blended learning. Blended learning emerged during the COVID-19 pandemic with the popularity of e-learning and ”combines online with offline”. Hybrid models can be beneficial for learning because they combine human interaction with the benefits of tech-driven personalization. This offers the potential to benefit from the best of both worlds. Reinforcement learning can guide the online part of learning and adjust content in real time. And the offline learning part can be adjusted according to the suggestions of the models, and live learning data can be collected. According to the author, this leads to increased autonomy and engagement and higher cognitive, social, and emotional aspects. The paper [Ingabire and Extension(2024)] also especially highlighted the increased flexibility in ”time, pace, and mode” of hybrid learning models.

## 6.2. Conclusion

The goal of this review was to showcase the potential and risks Reinforcement Learning has in e-learning systems. It was shown that Reinforcement Learning has already been successfully integrated in deployed e-learning systems, and millions of users already interact with it knowingly or unknowingly. Use cases range from curriculum sequencing to giving adaptive feedback, dynamically adjusting the difficulty, and spaced repetition. While the potential is immense, the loss of control over its own data is also an intriguing problem, and should not be forgotten. Future implementations in education should focus on hybrid learning. A promising idea is blended learning, which was described in 6.1. It combines offline with online learning and tries to optimize offline learning with online algorithms. A workflow in blended learning can be seen here 3. In my opinion, Reinforcement Learning does offer a powerful framework for developing smarter learning systems, which focus more on the single person and can give more precise learning instructions, but its success will depend on its integration into educational theory and the human part of learning, because without total human interaction other factors might be reduced.

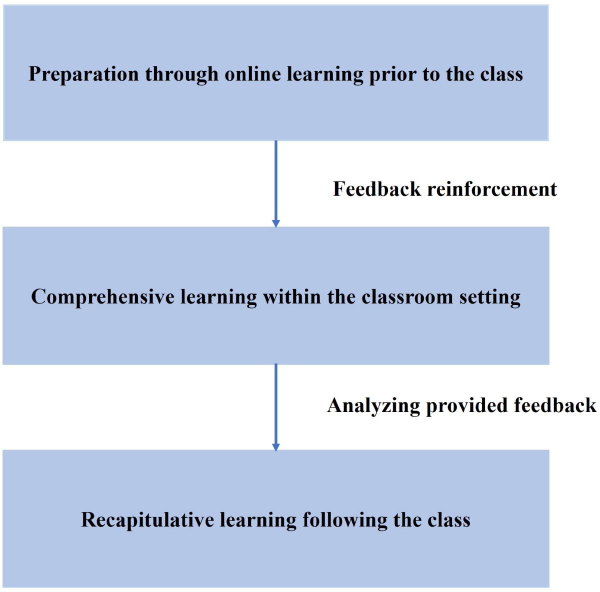


**Figure 3** The workflow in blended learning, taken from [Mon et al.(2023)Mon, Wasfi, Hayajneh, Slim, and Abu Ali].

# Bibliography

[Azhar et al.(2022)Azhar, Segal, and Gal] {Aqil Zainal} Azhar, Avi Segal, and Kobi Gal. Optimizing representations and policies for question sequencing using reinforcement learning. In *Proceedings of the 15th International Conference on Educational Data Mining, EDM 2022*, Proceedings of the 15th International Conference on Educational Data Mining, EDM 2022, United States, January 2022. International Educational Data Mining Society. doi: 10.5281/zenodo.6853123. Publisher Copyright: © 2022 Copyright is held by the author(s).; 15th International Conference on Educational Data Mining, EDM 2022 ; Conference date: 24-07-2022 Through 27-07-2022.

[Bartneck et al.(2021)Bartneck, Lütge, Wagner, and Welsh] Christoph Bartneck, Christoph Lütge, Alan Wagner, and Sean Welsh. *Privacy Issues of AI*, pages 61–70. 01 2021. ISBN 978-3-030-51109-8. doi: 10.1007/978-3-030-51110-4_8.

[Bicknell et al.(2023)Bicknell, Brust, and Settles] Klinton Bicknell, Chris Brust, and Burr Settles. How duolingo's ai learns what you need to learn. *IEEE Spectrum*, February 2023. URL `https://spectrum.ieee.org/duolingo`. Accessed: 2025-07-14.

[Chi et al.(2011)Chi, Vanlehn, Litman, and Jordan] Min Chi, Kurt Vanlehn, Diane Litman, and Pamela Jordan. Empirically evaluating the application of reinforcement learning to the induction of effective and adaptive pedagogical strategies. *User Model. User-Adapt. Interact.*, 21:137–180, 04 2011. doi: https://doi.org/10.1007/s11257-010-9093-1.

[Dake(2023)] Delali Dake. Reinforcement learning in education 4.0: Open applications and deployment challenges. *International Journal of Computer Science and Information Technology*, 15:47–61, 06 2023. doi: 10.5121/ijcsit.2023.15304.

[du Plooy et al.(2024)du Plooy, Casteleijn, and Franzsen] Eileen du Plooy, Daleen Casteleijn, and Denise Franzsen. Personalized adaptive learning in higher education: A scoping review of key characteristics and impact on academic performance and engagement. *Heliyon*, 10(21):e39630, 2024. ISSN 2405-8440. doi: https://doi.org/10.1016/j.heliyon.2024.e39630. URL `https://www.sciencedirect.com/science/article/pii/S2405844024156617`.

[Fahid and Chi(2021)] Ogan Amy Fahid, Abdul A. and Michelene T. H. Chi. Adaptively scaffolding cognitive engagement with batch constrained deep q-networks. In *Artificial Intelligence in Education (AIED 2021)*, volume 12748 of *Lecture Notes in Computer Science*, pages 280–292. Springer, 2021. doi: 10.1007/978-3-030-77857-6_22.

[Ghasemi and Ebrahimi(2024)] Majid Ghasemi and Dariush Ebrahimi. Introduction to reinforcement learning, 2024. URL `https://arxiv.org/abs/2408.07712`.

[Guskey(2007)] Thomas Guskey. Closing achievement gaps: Revisiting benjamin s. bloom's "learning for mastery". *Journal of Advanced Academics*, 19:8–31, 11 2007. doi: 10.4219/jaa-2007-704.

[Ingabire and Extension(2024)] Habimana Ingabire and Kiu Publication Extension. Hybrid learning models: Combining in-person and online education effectively. 3:16–19, 08 2024.

[Ma(2023)] Zhijuan Ma. Hybrid learning: A new learning model that connects online and offline. *Journal of Education and Educational Research*, 3:130–132, 05 2023. doi: 10.54097/jeer.v3i2.9059.

[Mon et al.(2023)Mon, Wasfi, Hayajneh, Slim, and Abu Ali] Bisni Fahad Mon, Asma Wasfi, Mohammad Hayajneh, Ahmad Slim, and Najah Abu Ali. Reinforcement learning in education: A literature review. *Informatics*, 10(3):74, 2023. doi: 10.3390/informatics10030074. URL `https://doi.org/10.3390/informatics10030074`.

[Paul(2024)] Joel Paul. Privacy and data security concerns in ai. 11 2024.

[Rafferty et al.(2011)Rafferty, Brunskill, Griffiths, and Shafto] Anna N. Rafferty, Emma Brunskill, Thomas L. Griffiths, and Patrick Shafto. Faster teaching by pomdp planning. In Gautam Biswas, Susan Bull, Judy Kay, and Antonija Mitrovic, editors, *Artificial Intelligence in Education*, pages 280–287, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg. ISBN 978-3-642-21869-9. doi: https://doi.org/10.1007/978-3-642-21869-9_37.

[Rahimi et al.(2023)Rahimi, Moradi, Vahabie, and Kebriaei] Masoud Rahimi, Hadi Moradi, Abdolhossein Vahabie, and Hamed Kebriaei. Continuous reinforcement learning-based dynamic difficulty adjustment in a visual working memory game, 08 2023.

[Sutton and Barto(2018)] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, 2nd edition, 2018. URL `http://incompleteideas.net/book/the-book-2nd.html`.

[Sweta(2021)] Soni Sweta. *Adaptive E-Learning System*, pages 13–24. 01 2021. ISBN 978-981-33-4680-2. doi: 10.1007/978-981-33-4681-9_2.

[Tam(2000)] Maureen Tam. Constructivism, instructional design, and technology: Implications for transforming distance learning. *Educational Technology Society*, 3, 04 2000.

[Yancey and Settles(2020)] Kevin P. Yancey and Burr Settles. A sleeping, recovering bandit algorithm for optimizing recurring notifications. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, KDD '20, page 3008–3016, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450379984. doi: 10.1145/3394486.3403351. URL `https://doi.org/10.1145/3394486.3403351`.

[Yulianti(2025)] Sonia Yulianti. The hidden bias in ai: How artificial intelligence reflects and reinforces social inequalities. 03 2025.