Gamification Framework for Reinforcement Learning-based Neuropsychology Experiments

Mounsif Chetitah
mounsif.chetitah@uni-wuerzburg.de
Julius-Maximilians University
Würzburg, Bavaria, Germany

Julian Müller
julian.mueller@stud.uni-wuerzburg.de
Julius-Maximilians University
Würzburg, Bavaria, Germany

Lorenz Deserno
Deserno_L@ukw.de
Universitätsklinikum
Würzburg, Bavaria, Germany

Maria Waltmann
Waltmann_M@ukw.de
Universitätsklinikum
Würzburg, Bavaria, Germany

Sebastian von Mammen
sebastian.von.mammen@uni-wuerzburg.de
Julius-Maximilians University
Würzburg, Bavaria, Germany

ABSTRACT

Reinforcement learning (RL) is an adaptive process where an agent relies on its experience to improve the outcome of its performance. It learns by taking actions to maximize its rewards, and by minimizing the gap between predicted and received rewards. In experimental neuropsychology, RL algorithms are used as a conceptual basis to account for several aspects of human motivation and cognition. A number of neuropsychological experiments, such as reversal learning, sequential decision-making, and go-no-go tasks, are required to validate the decisive RL algorithms. The experiments are conducted in digital environments and are comprised of numerous trials that lead to participants' frustration and fatigue. This paper presents a gamification framework for reinforcement-based neuropsychology experiments that aims to increase participant engagement and provide them with appropriate testing environments.

CCS CONCEPTS

• **Software and its engineering → Open source model**.

KEYWORDS

neuropsychology, gamification, serious games, reinforcement learning

ACM Reference Format:
Mounsif Chetitah, Julian Müller, Lorenz Deserno, Maria Waltmann, and Sebastian von Mammen. 2023. Gamification Framework for Reinforcement Learning-based Neuropsychology Experiments. In *Foundations of Digital Games 2023 (FDG 2023), April 12–14, 2023, Lisbon, Portugal.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3582437.3587190

FDG 2023, April 12–14, 2023, Lisbon, Portugal

ACM ISBN 978-1-4503-9855-8/23/04…$15.00
https://doi.org/10.1145/3582437.3587190

1 INTRODUCTION

RL is a process where learners or agents are trained to make a sequence of decisions in a given environment [9]. The learner does not receive instructions on how to make decisions but rather proceeds by trial and error. As a result of an action, the environment's state changes which gives rise to a reward, which, in turn, reinforces the performed behaviour [11].

In experimental neuropsychology, a field that studies human cognition and behavior, researchers use empirical methods to investigate the physiological structures of the nervous system and their effects on cognition and behavior. A variety of techniques can be used including laboratory research, neuropsychological tests, brain imaging, electroencephalography, and qualitative analysis in order to investigate human cognition. Since RL algorithms account for several aspects of cognition and motivation, including goal-directed control [2] and approach-avoidance behavior [1], they are considered a formal model for experimental neuropsychology [10]. These algorithms still need to be validated by conducting experiments, including reversal learning, sequential decision-making, and go-no-go tasks to measure patients' neural and physiological activity.

A sequential decision-making experiment involves participants choosing symbols from a screen within a specified time frame. The cycle repeats with new cues and actions, and experiments often involve multiple cycles. Deserno et al. [4] conducted a two-stage sequential decision-making task with 200 cycles. During the experiment, data is collected and analysed to understand human decision-making. Such experiments are conducted using digital environments to show the cues and enable the participant to perform actions. Considering that an experiment is composed of several cycles, motivating participants is usually achieved through monetary rewards. This means experimenters generally do not take into consideration the suitability of the testing environment. Yet, given the nature of the experiments, any support towards participatory engagement, effective study designs and the retrieval of meaningful results is beneficial.

In this paper, we present a gamification framework for RL-based experiments that supports these three goals. To the best of our knowledge, no such framework has previously been published. Our framework was developed by an interdisciplinary team of neuropsychologists and computer scientists using a user-centered approach. We elaborate more about this process and its conceptual results in Section 3, but not before providing some scientific background for our work in Section 2.

2 BACKGROUND

Gamification has been applied in many fields and industries [5], and many frameworks have been proposed to guide the design and implementation of gamified systems [7]. In the context of RL-based neuropsychology experiments, many applications have been produced. A two-stage RL task was used, for instance, to observe how humans' control approach changes as they grow up from model-free control (where actions are carried out based on the outcomes and strategies are adjusted for optimal rewards) to model-based control (when the agent attempts to understand its environment and create a model of it based on its interactions with it). [3]. The gamification relied on a cover story about space exploration. Participants can choose one of two spaceships or one of two alien characters in the first choice. If the player's selections result in a reward, a "space treasure" will appear. Influenca is a one-state RL task with two actions and win probabilities that change over time using Gaussian random walks [8]. The gamification relies on a story on combating viruses and selecting between different medications. The changing probabilities are explained with the viruses adapting to the used medication. Additionally, scores and completion rewards in the form of better lab equipment (a different background image) are used to motivate the players. After playing ten levels for free, players have to complete a questionnaire to unlock the remaining 20 levels. 127 players who completed 2904 levels were used to quantify intra- and inter-individual differences in value-based decision-making over time, demonstrating the game and its customizable app can be used for repeat assessment of reward learning.

These games combine RL elements with game design elements aimed at engaging the player as well as providing the basis for the experiments. Our paper synthesizes the previous work and proposes a systematic approach to the design of reinforcement-based experiments, beginning with the description of the experiments and progressing to software design.

3 FRAMEWORK

An RL-based, gamified experiment can be designed considering the following four elements: (1) A description of the experiment, (2) an RL model, (3) a game design document, and (4) a data collection module. An event-based software architecture integrates elements (1) to (4). First, we describe the experiment in order to identify key parameters (time constraints, the number of trials, and the number of stages per trial, etc.). Next, we describe the RL environment model using cues, states, rewards, and transition probabilities. As a third step, we add a layer of game design elements (mechanics, aesthetics, story, etc.). Finally, we determine which data will be collected. We provide a programming framework to address these four steps and which integrates the respective definitions by means of an event-based architecture. The resulting software artefact serves as boilerplate code that can be easily adjusted and used by designers and developers to support neuropsychologists.

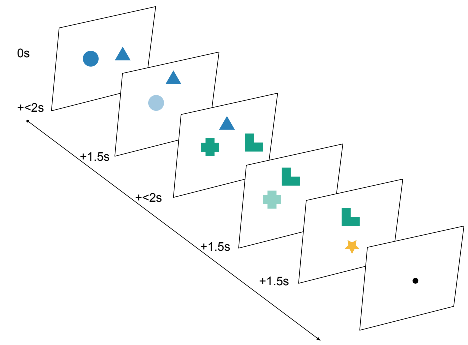


Figure 1: Example of subsequent displays in a sequential decision-making task as used in [4].

Table 1: Sequential Decision-Making Task Parameters

Parameter	SDM-2ST
Start state	First Choice
Number of states (including start state)	3
Transition probabilities	0.3 and 0.7
Number of trials	200
Number of stages per trial	2
Inter-trial time	2.0s
Decision time	1.5s
Choice highlight time	1.5s
Reward highlight time	1.5s

3.1 Experiment Description

In he sequential decision-making 2-stage task (SDM-2ST) from [4], participants make decisions in two stages across 200 trials, with a potential reward at the end. Participants have 2 seconds to select between two cues per stage, with cues randomly placed on the screen. If no response is made, the trial is terminated. After the decision phase, the chosen cue is highlighted for 1.5 seconds. See Figure 1 for an example cycle. The inter-trial period follows an exponential distribution with an average of 2 seconds. Details are summarized in Table 1.

3.2 RL Environment

The RL environment model used in the sequential decision-making experiment requires knowledge of states, transition probabilities, and rewards. Figure 2 illustrates a simplified version of the model, where the participant selects a cue in the first choice state and transitions to the second choice state based on the chosen cue. Transitions are probability-based, with all probabilities summing to 1. There are no rewards for transitions between the first and second choice states, and four transitions lead to a new state, with one cue having a 70% win probability and the other three having a 30% win probability.

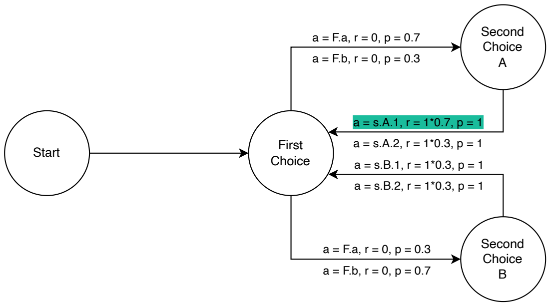


Figure 2: Simplified environment model for the two stage decision task

Table 2: Sequential Decision-Making Task Parameters Matched with Game Design Elements

Parameter	RL-model	Game Design
Number of trials	-	Game duration
Decision time	-	Time Pressure
Choice highlight time	-	Player Input
Reward highlight time	-	Reward
-	Cues	Game entities
-	Transitions	Mechanic
-	Rewards	Score system

3.3 Game Design

In the game design step, we associate elements from the experiment and the RL environment model with elements of game design. An experiment can use decision time as a time pressure mechanism to assist the participant in focusing on the task at hand, cues can be converted into game objects, etc.

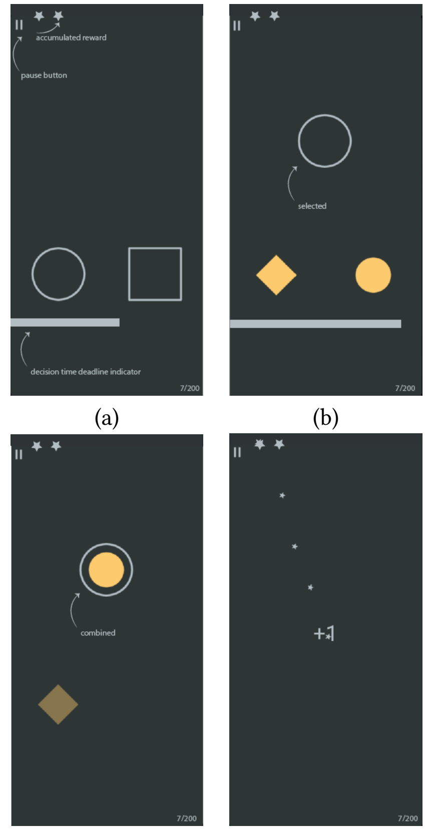
As for game design, the only constraint is ensuring that game elements are associated with RL elements and using the experiment parameters (see Table 2). Aside from that, the game's mechanics, dynamics, aesthetics, and story can all be tailored to the designer's preferences.

Figure 3 displays Clickn'Win, a mobile game demonstrating the framework. Besides the cumulative reward, it shows the previously selected cues below it. In the bottom-right corner, a progress bar indicates remaining decision time. To reduce cognitive load, the first and second cues are combined to form a complex shape in Step One, as shown in Figure 3 (c). Goal is to choose the highest-paying combination.

3.4 Data Collection

Data collection is available as a module in our framework (see data sample in Figure 4). The module considers the following data items:

(1) The trial number,
(2) a timestamp before the trial (when the inter-trial screen is shown),
(3) a timestamp at the start of the trial (when the first decision is shown),



(a) (b) (c) (d)

Figure 3: The game Clickn'Win reduces the memorization load by showing the selected cues from stages one (a) and two (b) in combination (c), before a reward is displayed (d).

A	C	G	H	I	J	M	N	O	P	Q
Trial Nr	Timestamp Before	First Choice Reacti	First	Common	Second Stage	Second Choice Re	Second	Win Prob	Reward	Timestamp End
4	09/30/2021 19:31:01	00:00:00.8858740	First A	yes	Second 1	00:00:01.5361860	Second 1 B	30.0%	no	09/30/2021 19:31:13
5	09/30/2021 19:31:13	00:00:01.2452290	First A	yes	Second 1	00:00:00.9310130	Second 1 A	30.0%	no	09/30/2021 19:31:19
6	09/30/2021 19:31:19	00:00:01.2220470	First B	no	Second 1	00:00:00.8300190	Second 1 B	30.0%	no	09/30/2021 19:31:27
7	09/30/2021 19:31:27	00:00:00.9753440	First A	no	Second 2	00:00:00.5938620	Second 2 A	30.0%	no	09/30/2021 19:31:36
8	09/30/2021 19:31:36	00:00:00.8267270	First A	yes	Second 1	00:00:00.9744670	Second 1 A	30.0%	yes	09/30/2021 19:31:44

Figure 4: A sample of the collected data

(4) a timestamp at the first decision, allowing to calculate the reaction time of the player,
(5) the first choice of the player, the resulting next state, and whether this transition was a common or rare transition,
(6) a timestamp each when the second decision is shown and when the second decision is made to calculate the reaction time in the second decision,
(7) the second choice of the player, its win probability and whether they actually received a reward,
(8) a timestamp at the end of the trial.

4 THE SOFTWARE ARCHITECTURE

We followed the model-view-controller software design pattern [6] for integrating the four outlined elements of our framework. Hereby, the RL environment model (Figure 5) serves as the model that connects four controllers with the application's view (Figure 6). The RL environment model is connected to sate, cue, reward, and transition objects. A state is composed of two cues that, when selected, initiate a transition to the next state. The reward has a value and a Boolean flag indicating whether it should be displayed to the user, or not. Transitions contain information about (1) the start state, (2) the selected cue, (3) the new state, (4) the reward, and (5) the probability of selection. In the given example, there are 3 state objects, 6 cue objects, 12 transition objects, 3 reward objects and 1 environment model object.

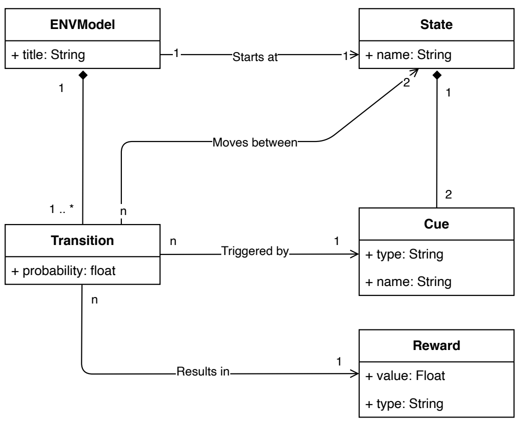


Figure 5: Class diagram representing all the entities that compose the RL environment model

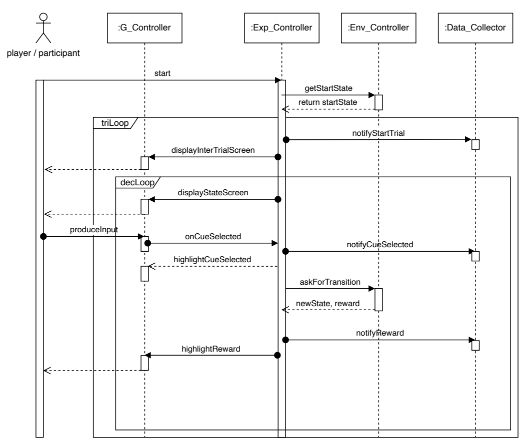


Figure 6: Sequence diagram representing all the events triggered and received by the controllers

A game controller, :G_Controller in Figure 6, runs the game loop of the application. It iterates through a set of trials and mediates between the player / participant and the experiment. The experiment controller, :Exp_Controller, ensures that the experiment is conducted as specified in accordance with the experiment's description as, for instance, specified in Table 1. As such, it is the time manager in the system to inform the player / participant about the concrete situation within the experiment and the logging module's :Data_Collection controller about timestamp-augmented interaction events. Finally, the environment controller, :Env_Controller, feeds the input data from the player / participant into the RL model as outlined above, which determines the software's reactions. We developed this architecture iteratively incorporating feedback from neuropsychologists, game designers, and computer scientists. We share the implementation in this repository [1].

5 CONCLUSION & FUTURE WORK

In this paper, we presented a modelling / programming framework for gamified neuropsychological RL experiments conducted by means of mobile devices. The framework requires its users to define four elements: (1) The description of the experiment, (2) an RL model, (3) a game design document, and (4) a data collection module. An according set of controller instances ensures that these respective perspectives are well-integrated. The resulting games are driven by a concrete instance of a generalized RL model. Based on the very concrete and established sequential decision-making task, we configured the framework and composed different gamified versions to implement this experiment. Our framework has been validated by an inter-disciplinary team of neuropsychologists and computer scientists and could certainly be used in more rigorous studies on a larger scale. To this end, the games could be enhanced in terms of game design and personalization to different participant groups. It could also be tested if adding more user interactions and mechanics that do not directly influence the experiment task and may therefore interfere with the data collection. Finally, developing more elaborate games like 3d endless runners would be thinkable. However, this would raise the question of whether a complex and action-packed game of that genre could work as a gamification of the two-stage decision task.

REFERENCES

[1] Charles S Carver. 2006. Approach, avoidance, and the self-regulation of affect and action. *Motivation and emotion* 30, 2 (2006), 105–110.
[2] Peter Dayan. 2009. Goal-directed control and its antipodes. *Neural Networks* 22, 3 (2009), 213–219.
[3] Johannes H Decker, A Ross Otto, Nathaniel D Daw, and Catherine A Hartley. 2016. From creatures of habit to goal-directed learners: Tracking the developmental emergence of model-based reinforcement learning. *Psychological science* 27, 6 (2016), 848–858.
[4] Lorenz Deserno, Quentin JM Huys, Rebecca Boehme, Ralph Buchert, Hans-Jochen Heinze, Anthony A Grace, Raymond J Dolan, Andreas Heinz, and Florian Schlagenhauf. 2015. Ventral striatal dopamine reflects behavioral and neural signatures of model-based control during sequential decision making. *Proceedings of the National Academy of Sciences* 112, 5 (2015), 1595–1600.
[5] Nadija Figol, Tetiana Faichuk, Iryna Pobidash, Olga Trishchuk, and Vasyl Teremko. 2021. Application fields of gamification. *Amazonia Investiga* 10, 37 (2021), 93–100.
[6] Adele Goldberg and David Robson. 1983. *Smalltalk-80: the language and its implementation*. Addison-Wesley Longman Publishing Co., Inc.
[7] Alberto Mora, Daniel Riera, Carina Gonzalez, and Joan Arnedo-Moreno. 2015. A literature review of gamification design frameworks. In *2015 7th international conference on games and virtual worlds for serious applications (VS-Games)*. IEEE, 1–8.
[8] Monja P Neuser, Franziska Kraeutlein, Anne Kuehnel, Vanessa Teckentrup, Jennifer Svaldi, and Nils B Kroemer. 2021. Influenca: a gamified assessment of value-based decision-making for longitudinal studies. *bioRxiv* (2021).
[9] Richard S Sutton and Andrew G Barto. 2018. *Reinforcement learning: An introduction*. MIT press.
[10] Richard S Sutton, Andrew G Barto, et al. 1999. Reinforcement learning. *Journal of Cognitive Neuroscience* 11, 1 (1999), 126–134.
[11] Marco A Wiering and Martijn Van Otterlo. 2012. Reinforcement learning. *Adaptation, learning, and optimization* 12, 3 (2012), 729.

[1]https://github.com/Julian24816/two-stage-task/tree/main