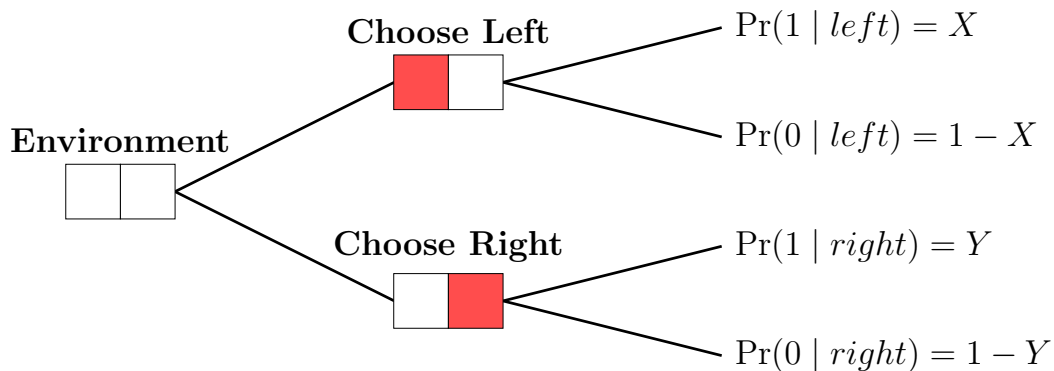# Welcome to the experiment!

We are interested in how humans make value based decisions and how these decisions are neurophysiologically grounded in the brain. The focus of this current study lies on a particular framework called *Decisions from experience* as illustrated by the following scenario.

> Before purchasing a new laptop, a consumer may look for information on the relative merits of various available models such as online reviews, prices, or technical specifications. We call this search for information a sequential sampling of the available options. At some point, a consumer must stop the search for information and choose one laptop model, often without having obtained full certainty about which model is the best.

Our main interest for the present study are the neural mechanisms that underlie the sequential allocation of active sampling efforts to the choice options in question. Furthermore, we are interested in how humans arrive at a decision to terminate their information search.

To study these mechanisms, we have devised a set of paradigms in which you have agreed to participate today. These paradigms represent the scenario described above in simple terms. You will always have to decide between two options that both represent a lottery. Each of these lotteries will have a probability to yield either a winning outcome or a losing outcome. However, these probabilities are unknown and have to be learned through sampling.

Choose Left $\quad \Pr(1 \mid left) = X$

$\Pr(0 \mid left) = 1 - X$

Environment

Choose Right $\quad \Pr(1 \mid right) = Y$

$\Pr(0 \mid right) = 1 - Y$

Figure 1: A binary lottery. X and Y refer to some probabilities.

Although all tasks follow this general principle, there will be small differences between them. We will describe each of the tasks in more detail right before it starts. Furthermore, you will always be allowed to practice the task before you start. If you have questions at any point, do not hesitate to ask the experimenter.
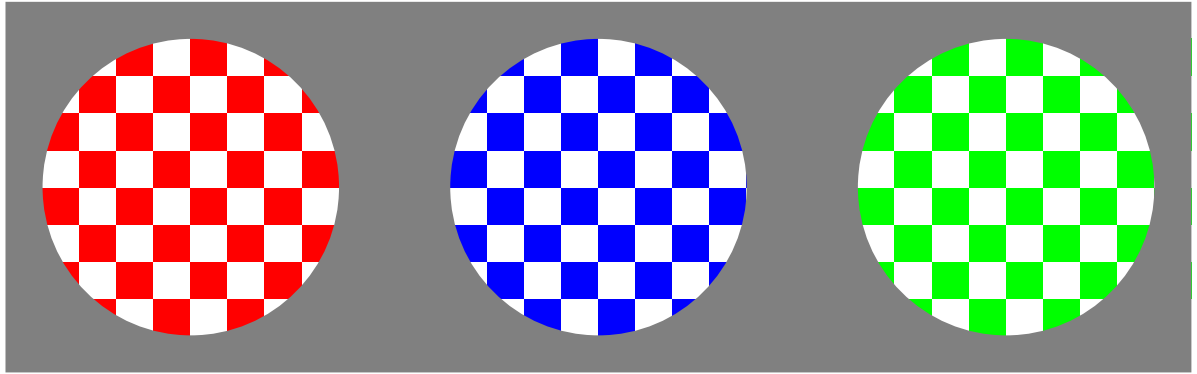
Figure 2: Three stimuli against a background as used in the experiment . The red and blue stimuli represent either a win or lose outcome. The green stimulus represents a distractor stimulus.
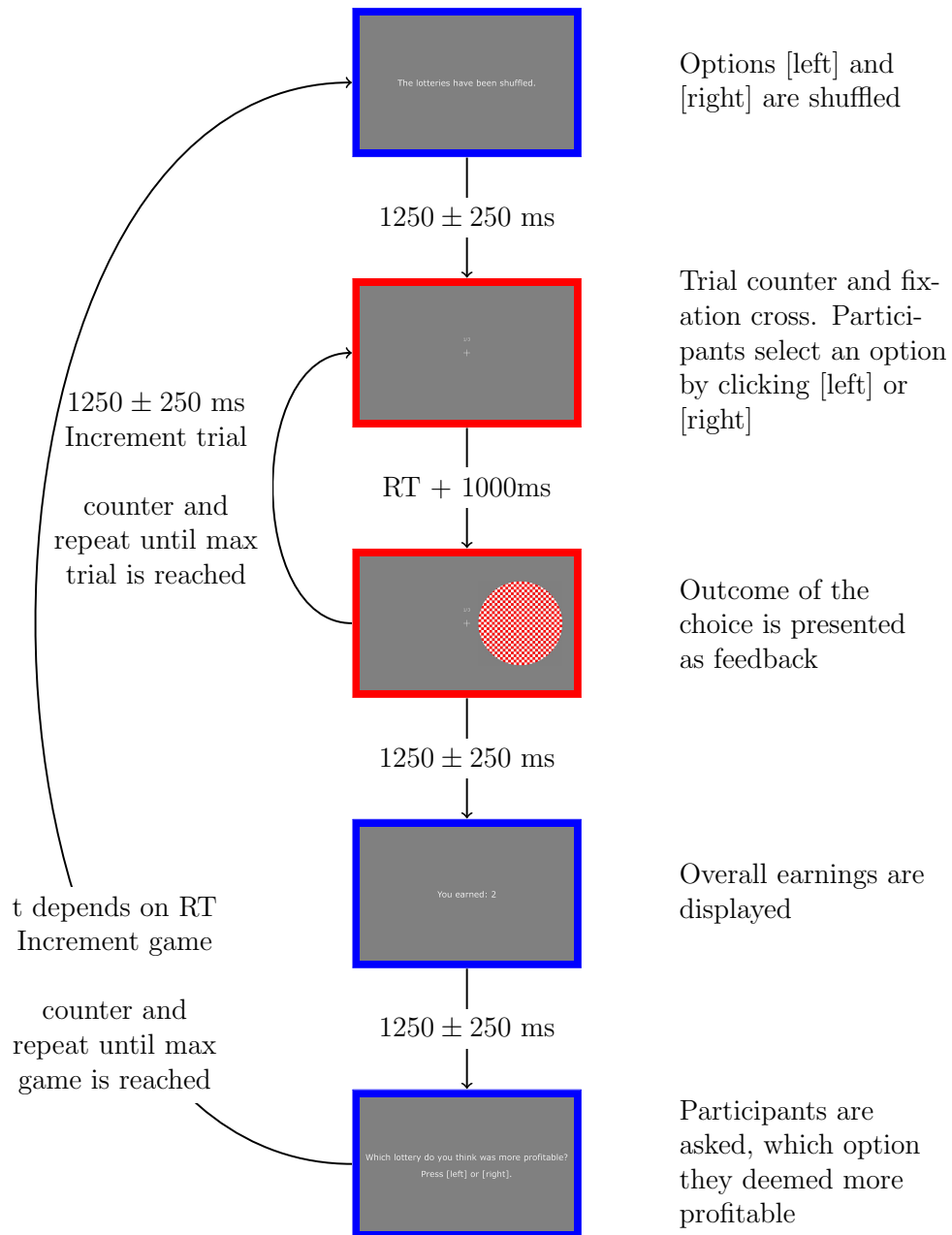
Bandit Paradigm

This task consists out of a structure, where we have 15 separate *games*, in each of which 20 *trials* are nested. For each trial, you must choose between two options [left] and [right]. These options represent binary lotteries, which have a certain probability to yield a winning outcome and a certain probability to yield a losing outcome. Your task is, to maximize the winning outcomes for each game that you play. This is most easily done by exploring the options a bit, and learning their underlying probabilities to yield winning or losing outcomes. However, it is important to realize that each outcome that you see contributes to your payoff.

Thus, it is a good strategy to try to balance exploration of the options and exploiting the knowledge you have already gained about the options.

You dont know the probabilities of the lotteries in advance, but you do know that during one game with all of its 20 trials, the probabilities of the lotteries and their locations [left] and [right] remain stable. Once all trials within one game have been completed, the lotteries are shuffled and a new game is started. Because there are 15 games, the lotteries will be shuffled 15 times and you will thus have 15 independent learning situations.
At the end of the overall experiment, we will calculate the percentage of winning outcomes to all outcomes and multiply it with a monetary amount, that you can earn in Euros additionally to your show up fee.

Figure 3: Experimental flow of the bandit paradigm. Red colors indicate the trial loop where participants explore and exploit the options. Once all trials have been spent within the trial loop, a transition to the blue game loop occurs, resetting the environment for a new game and leading to the trial loop again, until all games are exhausted. Note: RT=Reaction time, ms=miliseconds, t=time

Passive Condition

This task is very similar to the Bandit task. However, instead of choosing between options and learning their underlying probabilities to yield winning or losing outcomes, you are watching a replay. Your task is to fixate upon the fixation cross and experience the replay of a Bandit task. Importantly, there will be one addition: Occasionally, a known outcome will be replaced by a distractor outcome (green). As soon as you realize that such a distractor trial has happened, you have to press the [space] button.

At the end of the overall experiment, we will calculate your reaction time to the distractor trials. From all of these reaction times, we will calculate the percentage that were below a 400 miliseconds threshold and we will multiply it with a monetary amount that you can earn in Euros additionally to your show up fee.
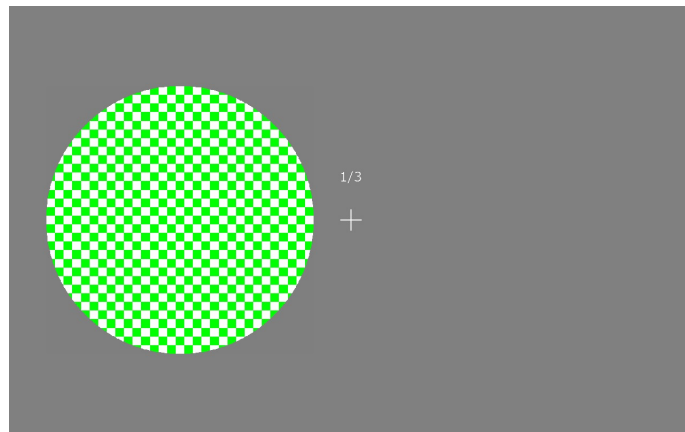
Figure 4: An outcome during a distractor trial.
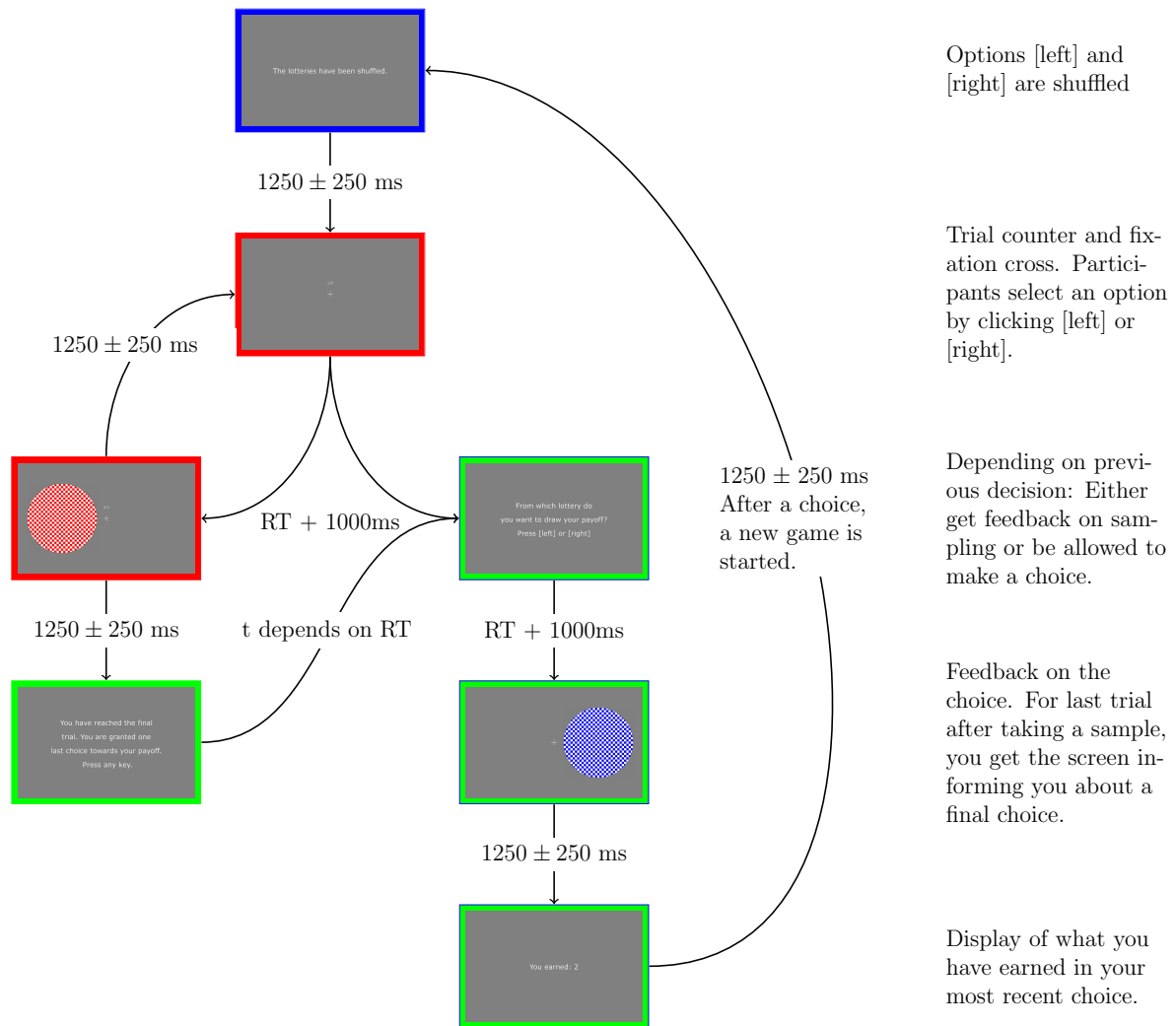
Sampling Paradigm

This task consists out of a structure, where we have separate *games*, in which *trials* are nested. For each trial, you can choose between two options [left] and [right]. These options represent binary lotteries, which have a certain probability to yield a winning outcome and a certain probability to yield a losing outcome.

However, you can also choose to not sample one of the options [left] or [right] and instead press [down]. When you press [down], you get to a stage of the game, where you must decide, which of the options [left] or [right] you find preferable. Then, an outcome will be generated from the option that you preferred. Your task is to maximize the winning outcomes that are obtained from this last game stage, after pressing [down].

Thus, the structure of the task is separated into two parts: During the first part, you can freely sample the options [left] and [right] to learn their underlying probabilities to yield winning or losing outcomes. After learning the options to a sufficient degree, you can exploit your knowledge of the options and press [down]. This will bring you to the stage of the game, where you can decide for one lottery over the other and draw your outcome that counts towards your payoff.

At the end of the overall experiment, we will calculate the percentage of winning outcomes in the final stage to all outcomes in the final stage and multiply it with a monetary amount, that you can earn in Euros additionally to your show up fee.

Overall you will be granted 300 trials in the exploration stage. The second stage (exploitation stage) does not count towards these trials. Note that a person who explores for 300 trials and makes only one choice in the second stage can still earn just as much money as a person who explores only for few trials and goes to the second stage very often.

Figure 5: Experimental flow of the sampling paradigm. Red colors indicate the sampling loop, from which one can transition to the blue choice loop, leading back to the sampling loop. Once the maximum of trials has been reached, a transition to the green loop for a last choice occurs. Note: RT=Reaction time, ms=miliseconds, t=time.