Initial experience has lasting effect on repeated risky decisions

Andrea Kóbor 1,+,*, Zsófia Kardoš 1,2,+, Ádám Takács 3, Noémi Éltető 4, Karolina Janacsek 5,6,7, Eszter Tóth-Fáber 8,5,6, Valéria Csépe 1,9, Dezso Nemeth 6,5,10,*

1: Brain Imaging Centre, Research Centre for Natural Sciences, Magyar tudósok körútja 2., H–1117, Budapest, Hungary
2: Department of Cognitive Science, Budapest University of Technology and Economics, Egy József utca 1., H-1111, Budapest, Hungary
3: Cognitive Neurophysiology, Department of Child and Adolescent Psychiatry, Faculty of Medicine of the TU Dresden, Fetscherstraße 74, 01307 Dresden, Germany
4: Max Planck Institute for Biological Cybernetics, Max-Planck-Ring 8, 72076 Tübingen, Germany
5: Institute of Psychology, ELTE Eötvös Loránd University, Izabella utca 46, H–1064, Budapest, Hungary
6: Brain, Memory and Language Research Group, Institute of Cognitive Neuroscience and Psychology, Research Centre for Natural Sciences, Magyar tudósok körútja 2., H–1117, Budapest, Hungary
7: Centre of Thinking and Learning, Institute for Lifecourse Development, School of Human Sciences, Faculty of Education, Health and Human Sciences, University of Greenwich, Old Royal Naval College, Park Row, 150 Dreadnought, SE10 9LS, London, United Kingdom
8: Doctoral School of Psychology, ELTE Eötvös Loránd University, Izabella utca 46, H–1064, Budapest, Hungary
9: Faculty of Modern Philology and Social Sciences, University of Pannonia, Egyetem utca 10, H–8200, Veszprém, Hungary
10: Lyon Neuroscience Research Center (CRNL), Université de Lyon, Centre Hospitalier Le Vinatier – Bâtiment 462 – Neurocampus 95 Boulevard Pinel 69675 Bron, Lyon, France

+: These authors contributed equally to this work.

*: Corresponding authors:
Andrea Kóbor
Brain Imaging Centre, Research Centre for Natural Sciences; Magyar tudósok körútja 2., H–1117, Budapest, Hungary.
E-mail: kobor.andrea@ttk.hu; Phone: +36 1 382 66 15.

Dezso Nemeth
CRNL – Lyon Neuroscience Research Center, Université Claude Bernard Lyon 1;
CRNL – Centre Hospitalier Le Vinatier – Bâtiment 462 – 95 Bd Pinel – 69675 Bron Cedex – Lyon, France.
E-mail: dezso.nemeth@univ-lyon1.fr; Phone: +33 4 81 10 65 46
Abstract

Primarily encountered information has been shown to influence perceptual and memory processes. However, it has been scarcely investigated how initial risk probabilities influence subsequent risk-taking behavior in an environment with repeated decisions. Therefore, the present study tested in two experiments whether young adults adjusted their choice behavior in the Balloon Analogue Risk task after an unsignaled and unexpected change point. The change point separated early trials from subsequent ones. While mostly positive (more reward) or mostly negative (no reward) events characterized early trials, subsequent trials were unbiased. In Experiment 1, the change point occurred after one-sixth or one-third of the trials without intermittence whereas in Experiment 2, it occurred between separate task phases. In Experiment 1, if negative events characterized the early trials, after the change point, risk-taking behavior increased as compared with the early trials. Conversely, if positive events characterized the early trials, risk-taking behavior decreased after the change point. Although participants shifted their choice behavior, the difference in risk taking due to initial experience remained, especially when this experience involved only one-sixth of the trials. In Experiment 2, slower adjustment of risk-taking behavior was observed after initially experiencing negative events than positive or baseline events. However, as all participants overcame the effect of initial experience, they became more prone to take risks. Altogether, results of both experiments not only suggest the profound effect of early experience but also that individuals dynamically update their risk estimates to adapt to the continuously changing environment.

Keywords: Balloon Analogue Risk Task (BART), behavior adjustment, decision making, initial luck, risk-taking behavior, prior experience
Introduction

Effective decision making is essential in healthy daily functioning. Effectiveness is supported by generating predictions about the probability of upcoming events. Many of these predictions are guided by inferences based on prior experiences with a similar situation or early experiences with the current situation. Considering risky choices, if we learned that regulations were moderate and had the repeated experience that driving beyond the speed limit did not end in negative consequences, we would continue this behavior more likely. However, in the case of stricter regulations, this habitual choice behavior becomes disadvantageous and should be overcome. Adaptation to the newly encountered regulations is probably gradual but takes an unknown amount of time, depending on several factors. Accordingly, this study investigates how persistent the effect of initial experience with risk probabilities is on subsequent risk-taking behavior after an unexpected and unsignaled change in these probabilities.

The lasting effect of primary experience has been shown in the fields of learning, memory, and decision making. One of the main questions has been for how long the already acquired knowledge remains persistent and how and when the updating of this knowledge happens. If the acquisition of probabilities underlying a certain task occurs in an incidental and unconscious manner and the change in these probabilities is unsignaled and unexpected, updating the already acquired knowledge might be hindered (e.g., 1,2–7). Meanwhile, some studies provided evidence also for the update of the already acquired implicit knowledge 8,9.

In the same vein, the investigation of experience-based decision making in nonstationary environments with risky choices and feedbacks on a trial-by-trial basis showed fast initial adaptation that was robust to subsequent changes, especially if explicit memory aids on previous outcomes were provided 10,11. Importantly, participants in these studies were made aware that outcome probabilities may change. Similarly, in more complex decision
settings (e.g., hypothetical stock market game), the initially established decision strategy was retained even though it became no longer optimal. In addition, a hint about the change, monetary incentives, or transfer to another task with the same deep structure but different surface could not adequately support participants in the update of their original decision strategy. Considering the previous studies together, it is conceivable that individuals would not adjust their sequential risky choices if the change in the decision environment were unsignaled and unexpected; however, this assumption should be directly tested.

Among tasks requiring sequential decisions, the Balloon Analogue Risk Task (BART) has been a widely used laboratory measure, because its structure and appearance mimics naturalistic risk taking. This task requires sequential decisions on risk in the form of inflating virtual balloons on a screen. Participants are told that each balloon pump is associated with either a reward or a balloon burst. Each successive pump not only increases the chance to gain more reward but also the probability of balloon burst and the accumulated reward to be lost. However, no information is provided to participants about the probability of balloon bursts. Therefore, they should learn the structure of the task in a trial-and-error manner to maximize the total reward as directed in the instructions.

It could be assumed that many studies using the BART and examining influential factors of risk taking on task performance (e.g., the effect of reward type, gender, age, situational factors, and clinical symptoms, see) could have involved the effect of initial experience in an uncontrolled manner. As the results of our explorative study presented in the Supplementary Material also indicate, the more successful pumps individuals experience before balloon bursts within the first five trials, the larger they tend to pump the balloons during the task (see also Supplementary Fig. S1). The persistent effect of initial experience possibly varies across individuals and might involuntarily contribute to the experimental
effects and group differences that have been the main questions of a given study. Overall, it seems that initial experience should be considered when using the BART.

Whether experience gathered on the early trials influences subsequent decision making in the BART has been scarcely investigated. The study of Koscielniak, et al. 23 showed in a sample of young and older women that initial good luck and bad luck (bursts on the first three balloons after 25, 27, and 28 pumps vs. bursts after three, four, and one pumps) modulated subsequent adaptation in a within-subjects design. While after the initially unlucky series, the number of balloon pumps gradually increased in the following phases of the task, after the initially lucky series, their number were higher only in the first phase following manipulation. However, the difference between the bad luck and good luck conditions remained persistent in every phase, showing lower risk taking in the former condition. Reanalysis of the same dataset revealed that, in both age groups, experiencing initial bad luck was associated with higher reward sensitivity, higher initial belief that the balloon will explode, and more uncertainty about this belief 24,25. The study of Bonini, et al. 26 used five early losses (bursts at the first, second, third, second, and first pumps) in a modified BART. They showed that pathological gamblers under treatment and healthy control participants reduced their risk-taking tendency after prior losses in this modified BART compared with the standard BART they also performed, while problem gamblers did not change their choice behavior. These studies manipulated only the very first trials by using extremely high or low balloon tolerances. However, smoother transitions in the payoff structure might be closer to real-world risky situations. In addition, longer initial experience might result in more persistent effect and weaker subsequent adaptation than shorter experience (cf. 10), but this issue has not been investigated. Thus, it remains to be tested whether the effect of initial experience is similarly profound if smoother transitions with different lengths are applied.
Accordingly, the present study directly tests how and to what degree initial experience influences risk-taking behavior in the BART by going beyond the previous studies \(^{23,26}\) in two aspects. As the first aspect, we do not manipulate “luckiness” with initial balloons tolerating extremely high or low pump numbers; instead, a relatively smooth transition in balloon tolerance is applied to change the payoff structure more gradually. As the second aspect, we test different lengths of initial experience or treat this experience as a separate initial phase of the task. In this way, we could more directly investigate how long the effect of initial experience lasts. Particularly, in Experiment 1, the first five (short manipulation) or ten (long manipulation) balloons of the task consisting of 30 balloons were manipulated. However, after Experiment 1, two questions remained open. First, it was not clarified whether early positive or negative experience had stronger influence on subsequent risk-taking behavior. Second, due to the different lengths of the manipulation, the underlying payoff structure differently changed across the experimental conditions, influencing the update of choice behavior intertwined with the effect of initial experience. Therefore, in Experiment 2, beyond positive and negative experience, we included baseline experience and separated the manipulated phase from the subsequent phase. The payoff structure of the subsequent phase that was used to quantify the update of choice behavior was the same across the experimental conditions. Originally, we assumed the profound and constraining effect of initial experience on risk-taking behavior in both experiments.

**Experiment 1**

In Experiment 1, using a modified BART, we manipulated the first five or ten balloons out of the 30 balloons. We created lucky and unlucky runs of trials by providing the possibility to inflate the balloons to a larger size (lucky condition) or to experience frequent balloon bursts even after relatively few balloon pumps (unlucky condition). We crossed these conditions
with the length of the initial experience (long condition: ten balloons vs. short condition: five balloons) to check whether the length of exposure to lucky or unlucky events differently influenced subsequent risk-taking behavior. These manipulations were based on the assumption that participants tended to gain some experience with the task in terms of action-outcome mapping (i.e., they tended to inflate the balloons instead of being risk-averse, \(^{18,27}\)), which would have enabled them to experience lucky and unlucky runs of events.

**Method**

**Participants.** Eighty healthy young adults took part in Experiment 1, 20 in each experimental condition. Participants were students from different universities in Budapest and young adult volunteers. All participants had normal or corrected-to-normal vision and none of them reported a history of any neurological and/or psychiatric condition. All of them provided written informed consent before enrollment. The experiment was approved by the United Ethical Review Committee for Research in Psychology (EPKEB) in Hungary and by the research ethics committee of Eötvös Loránd University, Budapest, Hungary; and, it was conducted in accordance with the Declaration of Helsinki. As all participants were volunteers, they were not compensated for their participation. Descriptive characteristics of participants randomly assigned to the four different experimental conditions in Experiment 1 are presented in Table 1.

**Stimuli, design, and procedure.** The appearance of the BART was the same as described in previous studies\(^ {19,20,28-30}\). This version of the task was written in Presentation (v. 18.1, Neurobehavioral Systems). According to the instructions, participants were asked to achieve as high score as possible by inflating empty virtual balloons on the screen. After each successful pump, the accumulated score on a given balloon (temporary bank) simultaneously increased with the size of the balloon. Instead of further pumping the balloon, participants
could have finished the actual balloon trial and collected the accumulated score, which was transferred to a virtual permanent bank. Two response keys on a keyboard were selected either to pump the balloon or to finish the trial. There were two possible outcomes as results of a pump: The size of the balloon together with the score inside increased (positive feedback) or the balloon burst (negative feedback). The balloon burst ended the actual trial, and the accumulated score on that balloon were lost, but this negative event did not decrease the score in the permanent bank.

One point was added to the temporary bank for the first successful pump, two for the second (i.e., the accumulated score for a given balloon was 3), three for the third (i.e., the accumulated score was 6), and so on. Five information chunks persistently appeared on the screen during the task: (1) the accumulated score for a given balloon in the middle of the balloon, (2) the label “Total score” representing the score in the permanent bank, (3) the label “Last balloon” representing the score collected from the previous balloon, (4) the response key option for pumping the balloon and (5) the other response option for collecting the accumulated score. After collecting the accumulated score and ending the balloon trial, a separate screen indicated the gained score. This screen or the other one indicating balloon burst was followed by the presentation of a new empty (small-sized) balloon indicating the beginning of the next trial.

Participants had to inflate 30 balloons in the task. However, in the four different conditions, the first five or ten balloons were manipulated (see Table 2). In the lucky long condition, the maximum number of successful pumps for each of the first ten balloons was between 12 and 15. In the lucky short condition, the maximum number of successful pumps for each of the first five balloons was 15. In the unlucky long condition, for each of the first ten balloons, this was between 4 and 5. In the unlucky short condition, for each of the first five balloons, this was 5. We also determined the maximum number of successful pumps for
each balloon in the remaining 20 or 25 balloons. These were identical to one another between the equal-length conditions (i.e., lucky vs. unlucky long conditions and lucky vs. unlucky short conditions, see Table 2). These values were a random sequence of integers generated from a uniform distribution between three and 15 in Matlab 8.5 (MathWorks Inc.). We have chosen this interval between minimum and maximum values to ensure that the remainder of the experiment was different from the initial phase and balloons could be, on average, inflated up to a “medium” size. There was no significant difference in the mean of the maximum number of successful pumps – calculated for the whole task – between the lucky long ($M = 10.66, SD = 3.44$) and lucky short ($M = 10.03, SD = 3.47$) conditions ($p = .412$) and between the unlucky long ($M = 7.60, SD = 3.28$) and unlucky short ($M = 8.36, SD = 3.05$) conditions ($p = .309$). Hence, the length of the initial phase did not change the overall potential of risk taking. Lucky long and unlucky long conditions differed significantly ($p < .001$) in the mean of the maximum number of successful pumps, so did lucky short and unlucky short conditions ($p = .023$).

Importantly, in all conditions, participants were not informed about the structure of the task and the change from the initial phase to the remaining phase in the maximum number of successful pumps. As we used a between-subjects design, one participant performed only one experimental condition. We would like to note that the terms “lucky” and “unlucky” were used to be consistent with the study of Koscielniak, et al. and to simplify the nomenclature of the design and the description of the results. As the subjective sense of luck or bad luck was not explicitly tested, no information was available on whether participants considered themselves as lucky or unlucky.

To test whether participants used any task-solving strategies and gained awareness about the regularities underlying the BART, which could have influenced their risk-taking behavior, we administered a short interview with two questions immediately after finishing
We asked participants (1) how they solved the task, how they tried to achieve as high score as possible; and (2) whether they have noticed any regularity in the sequence of balloon bursts. Participants’ answers were evaluated for the following aspects: (1) mentioning change between task phases (“Switching”), (2) mentioning any systematic task-solving approach throughout the task, irrespective of its adaptiveness (“Systematic”). For details and examples of these categories, see Supplementary Table S1. The distribution of the presence/absence of the two strategies are shown in Table 1. Note that post-experiment interviews did not target specifically these categories; therefore, we analyzed whether any of the categories dominated participants’ responses. This way, a participant’s response could have been assigned to more than one category.

**Data analyses.** Two dependent variables were calculated: the mean adjusted number of pumps across balloons (MAP; mean number of pumps on balloons that did not burst) and the percentage of balloon bursts. The MAP is a conventionally used index to measure deliberate, unbiased risk-taking behavior. The percentage (or number) of balloon bursts indicates the level of risk taking but also the effect of negative feedback throughout the task. Higher percentage of balloon bursts could also mirror higher propensity to test the structure of the task. However, because of the maximum number of pumps on each balloon is a priori fixed, this measure could be confounded, especially if the mean tolerance of a given run of balloons is low (see Fig. 1 and Fig. 3). To ease meta-analytic work, Table 1 shows overall performance on the different versions of the BART presented in this paper measured by these indices.

To check how long the effect of initial experience would persist, both the MAP and the percentage of balloon bursts were determined for five-balloon-long runs of the task; thus, we considered altogether six consecutive time bins of the task. Balloon bursts were expressed as the percentage of balloons that exploded in the given bin (i.e., if all five balloons exploded,
In the case of the long-manipulation conditions (lucky long and unlucky long conditions), the remaining phase of the task consisted of four bins (4 * 5 balloons). In the case of the short-manipulation conditions (lucky short and unlucky short conditions), the remaining phase consisted of five bins (5 * 5 balloons).

To evaluate the effect of initial experience on risk-taking behavior, first, three-way mixed analyses of variance (ANOVAs) were performed with Bin (1–6) as a within-subjects factor and Luck (lucky vs. unlucky) and Length (long vs. short) as between-subjects factors on the MAP and the percentage of balloon bursts. Second, on these dependent variables, two-way mixed ANOVAs with Bin and Luck as factors were performed separately in the equal-length conditions, i.e., for the lucky and unlucky long conditions and for the lucky and unlucky short conditions, respectively. We intentionally separated the equal-length conditions as the mean tolerance of balloons was comparable across the conditions only in four bins (i.e., the mean tolerance was the same in the following bins: Bin3_long = Bin2_short, Bin4_long = Bin3_short, Bin5_long = Bin4_short, Bin6_long = Bin5_short, see Table 2 and Fig. 1). Therefore, the possibility to inflate a balloon during the remaining phase was differently constrained across the conditions. Altogether, the second separate analysis could more clearly indicate the temporal change in risk-taking behavior after the initial manipulation has ended.

In all ANOVAs (here and in Experiment 2), the Greenhouse-Geisser epsilon (\(\varepsilon\)) correction\(^{33}\) was used when necessary. Original df values and corrected (if applicable) \(p\) values are reported together with partial eta-squared (\(\eta_p^2\)) as the measure of effect size. LSD (Least Significant Difference) tests for pairwise comparisons were used. Note that the numbers of pumps on balloons that did not burst were averaged across each time bin (five balloons) to obtain the MAP values that entered the ANOVAs. If a participant experienced only balloon bursts in a given bin, the corresponding MAP value was missing. Therefore, the
number of datapoints, the related df values, and the exact estimated marginal means varied across the analyses of the MAP in Experiments 1–2.

To compare the distribution of the presence of strategies across conditions, we conducted chi-squared tests. Exact significance tests were selected for Pearson’s chi-square, if the assumptions for a chi-squared test (at least 80% of the expected counts are more than five and all expected counts exceed one) were not met.
Table 1. Descriptive data of demographic variables and BART performance in Experiments 1–2 and in an explorative study presented in the Supplementary Material.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Explorative study</th>
<th>Experiment 1</th>
<th>Experiment 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition</td>
<td>–</td>
<td>Lucky long</td>
<td>Lucky short</td>
</tr>
<tr>
<td>n</td>
<td>156</td>
<td>20 (M)</td>
<td>20 (M)</td>
</tr>
<tr>
<td>Gender [Male/Female]</td>
<td>26/130</td>
<td>11 (M)</td>
<td>9 (M)</td>
</tr>
<tr>
<td>Age [years]</td>
<td>21.4 (3.9)</td>
<td>22.4 (1.6)</td>
<td>21.0 (1.1)</td>
</tr>
<tr>
<td>Education [years]</td>
<td>14.6 (2.1)</td>
<td>15.5 (1.3)</td>
<td>14.2 (0.9)</td>
</tr>
<tr>
<td>Mean adjusted pumps</td>
<td>8.3 (2.3)</td>
<td>8.7 (1.7)</td>
<td>9.3 (1.3)</td>
</tr>
<tr>
<td>Number of balloon bursts</td>
<td>12.4 (4.1)</td>
<td>11.3 (5.3)</td>
<td>14.2 (4.3)</td>
</tr>
<tr>
<td>Total score</td>
<td>711.7 (256.4)</td>
<td>797.2 (146.4)</td>
<td>758.6 (127.1)</td>
</tr>
<tr>
<td>Presence of strategies</td>
<td>–</td>
<td>no/yes</td>
<td>no/yes</td>
</tr>
<tr>
<td>Switching [%]</td>
<td>–</td>
<td>57.9/42.1</td>
<td>85.0/15.0</td>
</tr>
<tr>
<td>Systematic [%]</td>
<td>–</td>
<td>52.6/47.4</td>
<td>45.0/55.0</td>
</tr>
</tbody>
</table>

Note. Total score is the sum of reward (points) in the end of the task (this measure was not used in the present study). In Experiment 1, BART variables were calculated for the whole task (30 balloons), which included the initial phase with experimental manipulation. In Experiment 2, BART variables were calculated for the subsequent phase (30 balloons). For strategies, see Experiment 1 and Table S1. Responses to the interview questions regarding the applied strategy are missing for one participant in the lucky long, unlucky long, and unlucky conditions, respectively, because of technical reasons.
Table 2. Maximum number of successful pumps for each balloon in each condition in Experiments 1–2.

<table>
<thead>
<tr>
<th>Bin</th>
<th>Balloon</th>
<th>Lucky long</th>
<th>Lucky short</th>
<th>Unlucky long</th>
<th>Unlucky short</th>
<th>Lucky</th>
<th>Unlucky</th>
<th>Baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>15</td>
<td>15</td>
<td>5</td>
<td>5</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>15</td>
<td>15</td>
<td>5</td>
<td>5</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>12</td>
<td>15</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>12</td>
<td>15</td>
<td>4</td>
<td>5</td>
<td>19</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>15</td>
<td>15</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>15</td>
<td>8</td>
<td>5</td>
<td>8</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>15</td>
<td>8</td>
<td>5</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>12</td>
<td>12</td>
<td>4</td>
<td>12</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>15</td>
<td>13</td>
<td>5</td>
<td>13</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>12</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td>8</td>
<td>9</td>
<td>8</td>
<td>9</td>
<td>18</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>18</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>12</td>
<td>11</td>
<td>12</td>
<td>11</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td>13</td>
<td>12</td>
<td>13</td>
<td>12</td>
<td>15</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td>5</td>
<td>12</td>
<td>5</td>
<td>12</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>9</td>
<td>6</td>
<td>9</td>
<td>6</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>17</td>
<td></td>
<td>8</td>
<td>11</td>
<td>8</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>18</td>
<td></td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>11</td>
<td>18</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>19</td>
<td></td>
<td>12</td>
<td>5</td>
<td>12</td>
<td>5</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td>12</td>
<td>4</td>
<td>12</td>
<td>4</td>
<td>19</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>21</td>
<td></td>
<td>6</td>
<td>9</td>
<td>6</td>
<td>9</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>22</td>
<td></td>
<td>11</td>
<td>15</td>
<td>11</td>
<td>15</td>
<td>14</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>23</td>
<td></td>
<td>11</td>
<td>7</td>
<td>11</td>
<td>7</td>
<td>13</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>24</td>
<td></td>
<td>5</td>
<td>10</td>
<td>5</td>
<td>10</td>
<td>11</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>25</td>
<td></td>
<td>4</td>
<td>6</td>
<td>4</td>
<td>6</td>
<td>14</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>26</td>
<td></td>
<td>9</td>
<td>12</td>
<td>9</td>
<td>12</td>
<td>13</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>27</td>
<td></td>
<td>15</td>
<td>6</td>
<td>15</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>28</td>
<td></td>
<td>7</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>29</td>
<td></td>
<td>10</td>
<td>11</td>
<td>10</td>
<td>11</td>
<td>19</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>30</td>
<td></td>
<td>6</td>
<td>7</td>
<td>6</td>
<td>7</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>
Note. Initial phases during which we controlled for lucky, unlucky, or baseline experience are **boldfaced**. In Experiment 1, the initial phase was part of the main task including altogether 30 balloons; in Experiment 2, the initial phase included 10 separate balloons that were followed by the subsequent phase with 30 balloons. Dashed horizontal lines denote separate time bins, each consisting of five consecutive balloons. Experiment 1 consisted of six bins; Experiment 2 consisted of altogether eight bins, two bins in the initial phase and six bins in the subsequent phase. Numbers in *italics* in the column labelled as “Bin” denote the sequential numbers of time bins; those in the column labelled as “Balloon” denote the sequential numbers of balloons. Maximum number of successful pumps means that if a participant inflated the given balloon one pump further, the balloon burst. Note that the remainder sequence of balloons was the same between lucky long and unlucky long conditions, and between lucky short and unlucky short conditions in Experiment 1. In addition, the first 20 balloons in the remainder sequence of all conditions were the same in Experiment 1. In Experiment 2, the subsequent phase was the same across all conditions.

**Figure 1. Balloon tolerances in Experiment 1.** Mean balloon tolerances (bar charts, left y-axes) plotted against the mean percentages of balloon bursts (line charts with white square symbols, right y-axes) in each time bin of Experiment 1 split by condition (lucky long, unlucky long, lucky short, unlucky short) are presented. Time bins with the same balloon tolerances across conditions are violet-colored. The first two bins of the long manipulation are dark green (A, lucky long) and light green (B, lucky short), respectively. The first bins of the short
manipulation are dark orange (C, lucky short) and light orange (D, unlucky short), respectively. The different colors indicate that these initial bins slightly differ in mean tolerances between the long and short manipulation conditions. The final bins of the lucky short and unlucky short conditions have the same mean tolerance that is different from the other values. Error bars denote standard error of mean. (The visibility of error bars could be poor because of their low value and the resolution of the scale.)

Results

MAP in all conditions and time bins. The Bin * Luck * Length ANOVA on the MAP contrasting all six consecutive bins yielded significant main effects of Bin, $F(5, 335) = 4.15, \varepsilon = .628, p = .006, \eta_p^2 = .058$, Luck, $F(1, 67) = 101.74, p < .001, \eta_p^2 = .603$, and Length, $F(1, 67) = 9.36, p = .003, \eta_p^2 = .123$. These effects were qualified by the significant Bin * Luck, $F(5, 335) = 35.57, \varepsilon = .628, p = .006, \eta_p^2 = .347$, and Bin * Luck * Length interactions, $F(5, 335) = 6.99, \varepsilon = .628, p < .001, \eta_p^2 = .094$, indicating that the influence of luck and length on the MAP, i.e., higher MAP with lucky experience and higher MAP with short manipulation, differed as the task progressed (see Fig. 2A, B). The Bin * Length, $F(5, 335) = 1.75, \varepsilon = .628, p = .155, \eta_p^2 = .025$, and Luck * Length, $F(1, 67) = 0.11, p = .743, \eta_p^2 = .002$, interactions were not significant.

This analysis allowed us to check the effectiveness of initial manipulation in terms of luckiness. In the long manipulation, Bin1 and Bin2 encompassed the manipulated phase: The MAP was significantly higher in the lucky long than in the unlucky long condition in both bins, especially in Bin2 (Bin1: 8.63 vs. 4.34, $p < .001$; Bin2: 10.63 vs. 4.36, $p < .001$). In the short manipulation, Bin1 encompassed the manipulated phase: The MAP was significantly higher in the lucky short than in the unlucky short condition in Bin1 (9.73 vs. 4.74, $p < .001$), where the between-conditions difference across the task was the largest.

MAP in the lucky long and unlucky long conditions. To follow up the significant Bin * Luck * Length interaction, we separately analyzed the equal-length conditions. The Bin
**Luck ANOVA** contrasting the *lucky long* and *unlucky long* conditions on the MAP of the *remaining four bins* (Bin3, Bin4, Bin5, Bin6) yielded a significant main effect of Luck, $F(1, 35) = 7.91, p = .008, \eta^2_p = .184$, and a tendency for Bin, $F(3, 105) = 2.57, p = .058, \eta^2_p = .068$. The significant Bin * Luck interaction, $F(3, 105) = 10.67, p < .001, \eta^2_p = .234$, indicated that the effect of long initial manipulation on the MAP changed over the remaining time bins, after the manipulation has ended (see Fig. 2A). According to pairwise comparisons, while the MAP was significantly higher in the lucky long than in the unlucky long condition in Bin3 (8.22 vs. 5.17, $p < .001$), it did not differ in Bin4, Bin5, and Bin6 (all $ps \geq .068$). Particularly, in the lucky long condition, the MAP decreased from Bin3 to Bin4 (8.22 vs. 7.43, $p = .043$), and it was also lower in Bin6 than in Bin3 (7.05 vs. 8.22, $p = .014$). In contrast, in the unlucky condition, the MAP abruptly increased from Bin3 to Bin4 (5.17 vs. 6.92, $p < .001$), then, it remained consistent with a slight decrease by the end of the task (Bin4 > Bin6: 6.92 vs. 6.28, $p = .048$; Bin3 < Bin5: 5.17 vs. 6.68, $p < .001$; Bin3 < Bin6: 5.17 vs. 6.29, $p = .012$; the other pairwise differences in both conditions were not significant).

**Interim summary.** In terms of the MAP, the striking difference due to the long manipulation between unlucky and lucky participants attenuated in the second half of the task after experience with five more balloons (i.e., Bin3). Participants in the unlucky long condition quickly increased pumping while those in the lucky long condition slightly decreased this behavior in the remainder of the task (see Fig. 2A).

**MAP in the lucky short and unlucky short conditions.** The Bin * Luck ANOVA contrasting the *lucky short* and *unlucky short* conditions on the MAP of the *remaining five bins* (Bin2, Bin3, Bin4, Bin5, Bin6) yielded a significant main effect of Luck, $F(1, 35) = 35.87, p < .001, \eta^2_p = .506$, while the main effect of Bin was not significant, $F(4, 140) = 1.72, p = .148, \eta^2_p = .047$. The significant Bin * Luck interaction, $F(4, 140) = 4.84, p = .001, \eta^2_p = .122$, suggested, again, that the effect of short initial manipulation on the MAP changed over
the remaining bins, after the manipulation has ended (see Fig. 2B). According to pairwise comparisons, the MAP was significantly higher in the lucky short than in the unlucky short condition in all bins (all \( p < .009 \)). In the lucky short condition, the MAP slightly decreased by the end of the task (Bin2 > Bin6: 9.42 vs. 8.23, \( p = .011 \); Bin3 > Bin6: 9.06 vs. 8.23, \( p = .023 \)), while in the unlucky condition, it increased from Bin2 to Bin3; after that, it remained consistent (Bin2 < Bin3: 6.18 vs. 7.12, \( p = .004 \); Bin2 < Bin4: 6.18 vs. 7.35, \( p = .003 \); the other pairwise differences in both conditions were not significant).

**Interim summary.** In terms of the MAP, short manipulation yielded consistently decreased pumping behavior in the unlucky condition as compared with the lucky one over the entire task. However, after experience with five more balloons (i.e., Bin2), this between-conditions difference started to decrease (Bin3: \( \eta_p^2 = .438 \), Bin6: \( \eta_p^2 = .177 \)), since participants in both conditions adjusted their behavior in an opposite manner (see Fig. 2B).

**Balloon bursts in all conditions and time bins.** The Bin * Luck * Length ANOVA on the percentage of balloon bursts contrasting all six consecutive bins yielded significant main effects of Bin, \( F(5, 380) = 22.25, p < .001, \eta_p^2 = .226 \) and Luck, \( F(1, 76) = 10.32, p = .002, \eta_p^2 = .120 \). These effects were qualified by the significant Bin * Luck, \( F(5, 380) = 86.24, p < .001, \eta_p^2 = .532 \), Bin * Length, \( F(5, 380) = 47.14, p < .001, \eta_p^2 = .383 \), and Bin * Luck * Length, \( F(5, 380) = 18.30, p < .001, \eta_p^2 = .194 \), interactions. The Bin * Luck * Length interaction indicated that the overall more frequent balloon bursts with lucky experience varied as a function of length manipulation as the task progressed (see Fig. 2C, D). The main effect of Length, \( F(1, 76) = 2.24, p = .140, \eta_p^2 = .028 \), and the Luck * Length interaction, \( F(1, 76) = 3.16, p = .079, \eta_p^2 = .040 \), were not significant.

Effectiveness of initial manipulation was also checked. In the long manipulation, the percentage of balloon bursts was lower in the lucky long than in the unlucky long condition in both Bin1 and Bin2 (Bin1: 8% vs. 63%, \( p < .001 \); Bin2: 27% vs. 43%, \( p = .014 \)). In the short
manipulation, the percentage of balloon bursts was lower in the lucky short than in the unlucky short condition in Bin1 (14% vs. 58%, $p < .001$). The largest between-conditions differences across the task emerged during initial manipulation for both the long and short conditions.

**Balloon bursts in the lucky long and unlucky long conditions.** To follow up the significant Bin * Luck * Length interaction, we separately analyzed the equal-length conditions. In the Bin * Luck ANOVA contrasting the lucky long and unlucky long conditions on the percentage of balloon bursts of the remaining four bins (Bin3, Bin4, Bin5, Bin6), all effects were significant: Bin, $F(3, 114) = 55.60, p < .001, \eta^2_p = .594$, Luck, $F(1, 38) = 21.32, p < .001, \eta^2_p = .359$, and Bin * Luck, $F(3, 114) = 11.19, p < .001, \eta^2_p = .228$. These results suggest that the effect of long initial manipulation on balloon bursts varied across the remaining bins (see Fig. 2C).

Pairwise comparisons showed that while balloon bursts were significantly more frequent in the lucky long than in the unlucky long condition in Bin3, Bin4, and Bin5 (all $ps \leq .039$), it only tended to differ in Bin6 ($p = .058$). In the lucky long condition, balloon bursts were less frequent in Bin4 and Bin6 and more frequent in Bin3 and Bin5 (Bin3: 55%, Bin4: 32%, Bin5: 64%, Bin6: 40%; Bin3 > Bin4, Bin6, Bin3 < Bin5, all $ps \leq .029$; Bin4 < Bin5, $p < .001$; Bin5 > Bin6, $p < .001$). In the unlucky long condition, balloon bursts remained relatively infrequent and similar after manipulation, in Bin3 and Bin4, then, they became frequent in Bin5 and moderated in Bin6 (Bin3: 11%, Bin4: 7%, Bin5: 52%, Bin6: 25%; Bin3 = Bin4, $p = .399$; Bin5 > all other bins, all $ps < .001$; Bin4 < Bin6, $p < .001$; Bin6 > Bin3, Bin4, both $ps \leq .005$; the other pairwise differences in both conditions were not significant).

**Interim summary.** In terms of balloon bursts, participants previously experiencing a long series of unlucky balloons showed cautious behavior that considerably changed after testing ten more balloons. Those who experienced a long series of lucky balloons were more
inclined to take further risks and test the structure of the task (see Fig. 2C). However, as the mean tolerance of the balloons changed in terms of the maximum number of pumps, the percentage of balloon bursts decreased in both conditions (i.e., from Bin5 to Bin6, see Fig. 1A, B).

**Balloon bursts in the lucky short and unlucky short conditions.** The Bin * Luck ANOVA contrasting the lucky short and unlucky short conditions on the percentage of balloon bursts of the remaining five bins (Bin2, Bin3, Bin4, Bin5, Bin6) yielded similar results. Significant Bin, $F(4, 152) = 48.75, p < .001, \eta^2_p = .562$, Luck, $F(1, 38) = 49.00, p < .001, \eta^2_p = .563$, and Bin * Luck, $F(4, 152) = 6.47, p < .001, \eta^2_p = .148$, effects emerged (see Fig. 2D).

Pairwise comparisons showed that balloon bursts were significantly more frequent in the lucky short than in the unlucky short condition in all bins (all $ps \leq .004$). In the lucky short condition, balloon bursts were less frequent in Bin3 and more frequent in Bin4 (Bin2: 60%, Bin3: 34%, Bin4: 72%, Bin5: 53%, Bin6: 50%; Bin3 < all other bins, all $ps \leq .003$; Bin4 > all other bins, all $ps \leq .008$). In the unlucky short condition, balloon bursts were infrequent in Bin2 and Bin3, then, they became frequent in Bin4 and moderated in Bin5 and Bin6 (Bin2: 15%, Bin3: 7%, Bin4: 57%, Bin5: 29%, Bin6: 30%; Bin3 < Bin4, Bin5, Bin6, all $ps < .001$; Bin4 > all other bins, all $ps < .001$; Bin5 > Bin2, $p = .002$; Bin6 > Bin2, $p = .014$; the other pairwise differences in both conditions were not significant).

**Interim summary.** The pattern of change in balloon bursts over the remaining bins was similar in the short manipulation conditions to that of the long manipulation conditions. Participants previously experiencing a short series of unlucky balloons showed cautious behavior that considerably changed after testing ten more balloons during the remainder phase. Meanwhile, participants who experienced a short series of lucky balloons were more prone to take further risks and test the structure of the task (see Fig. 2D). However, as the
mean tolerance of the balloons changed in terms of the maximum number of pumps, the percentage of balloon bursts decreased in both conditions (i.e., from Bin4 to Bin5 and Bin6, see Fig. 1C, D).

**Task-solving strategies and awareness about regularities.** Participants in the lucky long condition gained more awareness of switching (“change in luck”) compared to those in the other conditions, $\chi^2(3) = 8.66, p = .033$. However, participants in the lucky long condition with awareness of switching did not differ from those without awareness of switching in the MAP and in the percentage of balloon bursts across the remaining bins (Bin3, Bin4, Bin5, Bin6), according to Bin * Awareness mixed ANOVAs (MAP: nonsignificant Awareness and Bin * Awareness effects, $F_s \leq 1.28, p \geq .278$; percentage of balloon bursts: nonsignificant Awareness and Bin * Awareness effects, $F_s \leq 2.07, p \geq .116$). There were no significant differences in the distribution of mentioning any systematic approach among conditions, $\chi^2(3) = 1.83, p = .638$ (see also Table 1). Altogether, it seems that any awareness about the change in the task structure did not alter the adjustment of behavior in the lucky long condition.

![Figure 2. Results of Experiment 1.](https://doi.org/10.1101/2020.09.25.313551)
respectively. Vertical yellow thin lines indicate that initial manipulation ended after the first (short manipulation) or second (long manipulation) time bin. Error bars denote standard error of mean. (The visibility of error bars could be poor because of their low value and the resolution of the scale.)

**Discussion**

In Experiment 1, we showed that initial experience influenced how participants adjusted their subsequent risk-taking behavior. Lucky compared with unlucky initial experience led to overall higher deliberate risk-taking behavior. More importantly, approximately five balloons after the change in structure, participants started to update their risk estimates, irrespective of initial experience: Pumping behavior increased following unlucky experience, while it decreased following lucky experience in the remainder of the task. The degree of this change differed as a function of how long participants experienced the manipulation: In the case of the long manipulation, pumping behavior of participants became similar in the second half of the task. Yet, there was a difference in the asymptotic behavior between participants experiencing lucky and unlucky events in the case of the short manipulation. Altogether, after initially unlucky events, they converged to a more risk-averse behavior in both length manipulation conditions (see Figure 2A, B).

Similarly, the percentage of balloon bursts reflected the overall more cautious behavior of participants initially experiencing unlucky events. However, the change in the percentage of balloon bursts also suggested that participants adjusted their behavior to the mean tolerance of balloons, irrespective of the length of manipulation and of experiencing lucky or unlucky events. Because of the different lengths of initial manipulation, the salient changes in the percentage of balloon bursts occurred five balloons later in the long conditions (Bin3, Bin5) than in the short conditions (Bin2, Bin4) but with the same dynamics. In this way, the length of the manipulated phase caused only an offset in risk-taking behavior if the balloon bursts were considered.
Although there was a difference between lucky and unlucky conditions in terms of overall risk-taking behavior and more subtle differences in how they adjusted this behavior in time, without quantifying the “baseline” effect of a neutral experience, this design did not enable us to determine whether lucky or unlucky experience was more influential. In addition, the number of balloons and their tolerance in the remaining phase of the task differed across conditions. Furthermore, participants in the lucky long condition became more aware of the fact that the task structure changed, although this awareness did not reliably alter their risk-taking behavior. Therefore, the second experiment aimed to address these shortcomings of Experiment 1.

**Experiment 2**

In Experiment 2, we went beyond Experiment 1 in two aspects. First, we manipulated a separate initial phase with ten balloons that preceded a subsequent phase with 30 balloons. Second, in this initial phase, we used three conditions: lucky, unlucky, and baseline. This design enabled us to analyze risk-taking behavior on the exact same sequence of 30 balloons across all conditions in the subsequent phase. Moreover, with the use of the baseline condition, we could test whether lucky or unlucky experience was more influential.

**Method**

**Participants.** Ninety young adults took part in Experiment 2, 30 in each experimental condition. None of them participated in Experiment 1. Participants were students from different universities in Budapest and young adult volunteers. All participants had normal or corrected-to-normal vision and none of them reported a history of any neurological and/or psychiatric condition. All of them provided written informed consent before enrollment. The experiment was approved by the United Ethical Review Committee for Research in
Psychology (EPKEB) in Hungary and by the research ethics committee of Eötvös Loránd University, Budapest, Hungary; and, it was conducted in accordance with the Declaration of Helsinki. As all participants were volunteers, they were not compensated for their participation. Descriptive characteristics of participants randomly assigned to the three different conditions are presented in Table 1.

**Stimuli, design, and procedure.** The appearance of the task was the same as in Experiment 1. Participants had to inflate altogether 40 balloons in this task. The first ten balloons belonged to the *initial phase*; the remaining 30 balloons belonged to the *subsequent phase*. We manipulated the balloons of the initial phase. In the *lucky* condition, we used the same sequence as in the lucky long condition of Experiment 1 in terms of the maximum number of successful pumps (12-15 pumps); similarly, in the *unlucky* condition, we used the same sequence as in the unlucky long condition of Experiment 1 (4-5 pumps). For the *baseline* condition, we generated a random sequence of ten integers from a uniform distribution between two and 19, as balloon burst was enabled only after the third pump, and the maximum number of successful pumps was 19 in those versions that were used in our previous studies (e.g., 28). We determined the maximum number of successful pumps for each of the 30 balloons in the subsequent phase, and these values were identical across the different conditions (see Table 2 and Fig. 3). These values were, again, generated from a uniform distribution between two and 19. The maximum pump values for each balloon in both phases, separately for each condition are presented in Table 2. The mean of the maximum number of successful pumps – calculated for the initial phase – significantly differed across the three conditions, $F(2, 29) = 16.53$, $p < .005$. It was significantly lower in the unlucky than in the baseline and lucky conditions (unlucky: $M = 4.60$, $SD = 0.52$, baseline: $M = 10.60$, $SD = 6.08$, lucky: $M = 13.80$, $SD = 1.55$; unlucky vs. baseline: $p = .001$, unlucky vs. lucky: $p < .001$); and, it tended to be higher in the lucky than in the baseline condition ($p = .059$). The mean of
the maximum number of successful pumps – calculated for the initial and the subsequent phases together (whole task) – did not differ significantly across the three conditions, $F(2, 119) = 2.00, p = .140$ (unlucky: $M = 8.88, SD = 5.26$, baseline: $M = 10.38, SD = 5.47$, lucky: $M = 11.18, SD = 4.92$; unlucky vs. baseline: $p = .201$; lucky vs. baseline: $p = .495$; unlucky vs. lucky: $p = .051$). Hence, the manipulation of the initial phase did not considerably change the overall potential of risk taking.

Participants were told that they were going to have ten balloon trials in the initial phase of the task and then 30 balloon trials in the subsequent phase. The starting score was zero in both phases, about which participants were also informed. Importantly, no information was provided about the change in the payoff structure. The short interview as described in Experiment 1 was also administered after the task.

**Data analyses.** As in Experiment 1, both the MAP and the percentage of balloon bursts were determined for five-balloon-long runs of the task. Altogether in the initial and subsequent phases, there were eight consecutive time bins in the task. First, to check the effectiveness of initial manipulation, two-way mixed ANOVAs were performed with Bin (1–2) as a within-subjects factor and Condition (lucky, baseline, unlucky) as a between-subjects factor on the MAP and the percentage of balloon bursts related to the initial phase. Second, to evaluate the effect of initial experience on subsequent risk-taking behavior, two-way mixed ANOVAs were performed with Bin (1–6) as a within-subjects factor and Condition (lucky, unlucky, baseline) as a between-subjects factor on the MAP and the percentage of balloon bursts related to the subsequent phase. For illustrative purposes, risk-taking behavior in the eight consecutive time bins of the entire task are presented (see Fig. 4). To compare strategies across conditions, we conducted chi-squared tests similarly to Experiment 1.
Figure 3. Balloon tolerances in Experiment 2. Mean balloon tolerances (bar charts, left y-axes) plotted against the mean percentages of balloon bursts (line charts with white square symbols, right y-axes) in each time bin of Experiment 2 split by condition (lucky, baseline, unlucky) are presented. Vertical yellow areas separate the initial and subsequent phases of the task, each containing two and six time bins, respectively. This short pause lasted a few seconds until the experimenter initiated the next phase of the task. Time bins with the same balloon tolerances across conditions are grey-colored. The first two bins of the lucky condition (A) are dark green as they were the same as in the lucky long condition of Experiment 1 (see Figure 1A). The first two bins of the unlucky condition (C) are light green as they were the same as in the unlucky long condition of Experiment 1 (see Figure 1B). These were different from the tolerance values of the first two bins of the baseline condition (B). Error bars denote standard error of mean. (The visibility of error bars could be poor because of their low value and the resolution of the scale.)

Results

MAP in the initial task phase. The Bin * Condition ANOVA on the MAP contrasting the two consecutive time bins of the initial phase yielded significant main effects of Bin, $F(1, 86) = 74.90, p < .001$, $\eta^2_p = .466$, and Condition, $F(2, 86) = 61.42, p < .001$, $\eta^2_p = $
The Bin * Condition interaction was also significant, $F(1, 86) = 53.58, p < .001$, $\eta^2_p = .588$. Pairwise comparisons showed that all conditions differed from one another in both bins (all $ps \leq .032$) with participants in the lucky condition showing the highest MAP (Bin1$_{initial}$: 7.60, Bin2$_{initial}$: 11.14), those in the unlucky condition showing the lowest MAP (Bin1$_{initial}$: 4.11, Bin2$_{initial}$: 4.49), and the baseline condition falling in-between (Bin1$_{initial}$: 6.64, Bin2$_{initial}$: 6.66). In addition, the difference between the two initial bins was significant only in the lucky condition ($p < .001$): These participants increased their pumping behavior during the initial phase (see Fig. 4A).

**MAP in the subsequent task phase.** From the Bin * Condition ANOVA on the MAP contrasting the six consecutive time bins of the subsequent phase, the main effect of Bin, $F(5, 420) = 80.62, \varepsilon = .563, p < .001$, $\eta^2_p = .490$, and the Bin * Condition interaction, $F(10, 420) = 3.00, \varepsilon = .563, p = .009$, $\eta^2_p = .067$, were significant, while the Condition main effect was not significant, $F(2, 84) = 1.12, p = .331$, $\eta^2_p = .026$. Pairwise comparisons showed that significant between-conditions differences emerged in Bin1 and Bin3. In Bin1, the MAP was lower in the unlucky than in the lucky (5.26 vs. 6.68, $p = .001$) and baseline (5.26 vs. 6.50, $p = .003$) conditions. In Bin3, the MAP was lower in the unlucky than in the lucky condition (6.62 vs. 7.97, $p < .005$).

Considering the change of MAP over time, it was similar in the baseline and lucky conditions. In the lucky condition, the MAP was the lowest in Bin2 (5.75; lower than in all other bins, all $ps < .001$), and it was also lower in Bin1 than in Bin3, Bin4, Bin5, and Bin6 (6.68, all $ps < .001$). The MAP abruptly increased from Bin2 to Bin3, then, it increased to a lesser degree and remained consistently high throughout the task (Bin3 < Bin4: 7.97 vs. 8.58, $p = .018$; Bin3 < Bin5: 7.87 vs. 8.85, $p = .005$). In the baseline condition, the MAP was the lowest, again, in Bin2 (5.70, lower than in all other bins, all $ps < .001$), and it was also lower in Bin1 than in Bin4, Bin5, and Bin6 (6.50, all $ps \leq .002$). A gradual increase occurred after
Bin2 throughout Bin3 (7.04, lower than Bin4, Bin5, Bin6, all $ps \leq .002$), Bin4 (7.90, lower than Bin5, $p = .006$), and Bin5 (8.53, higher than all the subsequent bins, all $ps \leq .006$); then, it did not further increase in Bin6. Meanwhile, the MAP was similarly low in Bin1 and Bin2 in the unlucky condition; then, it steeply increased until Bin5 and remained similar in Bin6 (Bin3: 6.62, Bin4: 8.06, Bin5: 8.94, Bin6: 9.16; all differences were significant, all $ps \leq .016$, except that of Bin5 vs. Bin6). (The other pairwise comparisons in this analysis were not significant.)

**Interim summary of the MAP results.** After competing ten balloons of the subsequent phase, all participants increased their risk-taking behavior in terms of the MAP. Therefore, the striking between-conditions differences of the initial phase equalized, especially in the second half of the task. It seems that participants in the lucky and unlucky conditions adjusted their risk-taking behavior to the changed payoff structure, as they showed more cautious (lucky condition) or riskier (unlucky condition) behavior across the first ten balloons of the subsequent phase as compared with the initial phase. Meanwhile, across the same ten balloons, those in the baseline condition continued their behavior already established in the initial phase (see Fig. 4A).

**Balloon bursts in the initial task phase.** The Bin * Condition ANOVA on the percentage of balloon bursts contrasting the two consecutive time bins of the initial phase yielded significant main effects of Bin, $F(1, 87) = 89.29$, $p < .001$, $\eta^2_p = .506$, and Condition, $F(2, 87) = 11.93$, $p < .001$, $\eta^2_p = .215$. The Bin * Condition interaction was also significant, $F(1, 86) = 29.18$, $p < .001$, $\eta^2_p = .401$. Pairwise comparisons showed more frequent balloon bursts in the unlucky than in the lucky (38.7% vs. 9.3%, $p < .001$) and baseline (38.7% vs. 14%, $p < .001$) conditions in Bin$_{1\text{initial}}$, while more frequent balloon bursts in the baseline than in the lucky (58% vs. 32%, $p < .001$) and unlucky (58% vs. 38.7%, $p < .001$) conditions in Bin$_{2\text{initial}}$. Participants experienced more balloon bursts in Bin$_{2\text{initial}}$ than in Bin$_{1\text{initial}}$ in both
the lucky (32% vs. 9.3%, $p < .001$) and baseline (58% vs. 14% vs. $p < .001$) conditions, while those in the unlucky condition showed virtually the same values (see Fig. 4B).

**Balloon bursts in the subsequent task phase.** From the Bin * Condition ANOVA on the percentage of balloon bursts contrasting *the six consecutive time bins of the subsequent phase*, the main effect of Bin, $F(5, 435) = 131.69$, $\varepsilon = .852$, $p < .001$, $\eta_p^2 = .602$, and the Bin * Condition interaction, $F(10, 420) = 8.87$, $\varepsilon = .852$, $p = .009$, $\eta_p^2 = .169$, were significant, while the Condition main effect was a tendency, $F(2, 87) = 3.08$, $p = .051$, $\eta_p^2 = .066$. Pairwise comparisons showed that significant between-conditions differences emerged in Bin1 and Bin3. In Bin1, the balloon bursts were the least frequent in the unlucky condition (28.6%), more frequent in the baseline condition (50%), and the most frequent in the lucky condition (61.3%); all conditions differed significantly from one another (all $p$s ≤ .003). In Bin3, balloon bursts were less frequent in the unlucky than in the lucky condition (28% vs. 37.3%, $p = .007$).

Considering the change in the frequency of balloon bursts, it was similar in the baseline and lucky conditions. In the lucky condition, balloon bursts were the most frequent in Bin1 (61.3%) and Bin6 (56%) and the least frequent in Bin4 (21.3%); although the frequency of balloon bursts decreased until Bin4, after that, it started to increase again (all pairwise differences, except Bin1 vs. Bin6 and Bin2 vs. Bin3, were significant, all $p$s ≤ .032). In the baseline condition, the frequency of balloon bursts gradually decreased from Bin1 (50%) until Bin4 (14.7%); after that, it increased again reaching the overall highest value in Bin6 (60%; all pairwise differences, except Bin2 vs. Bin3, were significant, all $p$s ≤ .011). Meanwhile, participants in the unlucky condition started the task with a lower frequency of balloon bursts that did not differ significantly across the first three bins (Bin1: 28.7%, Bin2: 32.6%, Bin3: 28%, all $p$s ≥ .151). As in the other conditions, the frequency of balloon bursts reached its
minimum in Bin4 (15.3%); then, it increased, showing the highest value in Bin6 (60%; all pairwise differences, except Bin5 vs. Bin1 and Bin5 vs. Bin3, were significant, all \( ps \leq .002 \)).

**Interim summary of the balloon bursts results.** In the subsequent phase of the task, participants in the lucky and baseline conditions rapidly updated their risk estimates in terms of decreasing the percentage of balloon bursts. This process required experience with five balloons at the beginning of the subsequent phase. Although participants in the baseline condition tested the payoff structure to the largest extent in the initial phase, they experienced less frequent balloon bursts than those in the lucky condition even within the first five balloons. Updating the risk estimates was slower in the unlucky condition: Their cautious behavior was reflected in the lower frequency of balloon bursts until Bin3 as compared with the other conditions. However, as all participants experienced a relatively low number of balloon bursts due to the high mean tolerance of the balloons in Bin3, Bin4, and Bin5 (see Fig. 3), they might have intended to pump the balloons to a “large” size in the end of the task (Bin6), contributing to the U-shaped pattern of the change (see Fig. 3, Fig. 4B). Possibly due to the combination of updated risk estimates and the change in mean balloon tolerance across bins, the slight between-conditions difference disappeared in the second half of the subsequent phase.

**Task-solving strategies and awareness about regularities.** There were no significant differences in the distribution of mentioning switch in luck between the two phases, \( \chi^2(1) = 2.59, p = .108 \), or in systematic approach, \( \chi^2(2) = 1.36, p = .569 \), across the experimental conditions (see also Table 1).
Figure 4. Results of Experiment 2. The group-average values of the MAP (A) and the percentage of balloon bursts (B) in each of the conditions (lucky, baseline, unlucky) are presented. Vertical yellow areas separate the initial and subsequent phases of the task, each containing two and six time bins, respectively. This short pause lasted a few seconds until the experimenter initiated the next phase of the task. Error bars denote standard error of mean. (The visibility of error bars could be poor because of their low value and the resolution of the scale.)

Discussion

Experiment 2 involved not only lucky and unlucky conditions but also a baseline condition; and, the initial phase with experimental manipulation was separated from the subsequent phase where the adjustment of risk-taking behavior was measured. Experience in the initial phase influenced risk-taking behavior in the subsequent phase. Participants in the baseline and lucky conditions quickly decreased risk taking in the subsequent phase; then, due to adaptation to the changed structure, they increased risk taking again. Participants in the unlucky condition became more risk takers in a slower manner; however, in the second half of the task, the modulating effect of initial experience across conditions was not observed.

Experience in the baseline condition yielded a behavioral pattern that fell between that of the unlucky and lucky conditions only in the initial phase (MAP) and in the first bin of the subsequent phase (balloon bursts). Otherwise, the behavioral pattern of the baseline condition was comparable to that of the lucky condition. Therefore, it is likely that unlucky experience
was more influential on later risk-taking behavior than lucky experience. Importantly, the entire sample became more prone to take risk by increasing the number of balloon pumps as the task progressed. The change in the experienced balloon bursts across bins also supports participants’ sensitivity to the underlying structure. According to the verbal reports, the experimental effects were not driven by strategy differences even though the pause between the two phases could have served as signal indicating an unspecific change.

**General Discussion**

**Summary of findings**

This study examined how initial experience with risk probabilities influenced later risk-taking behavior in a task involving sequential decisions. To this end, we inserted an unsignaled and unexpected change point in the payoff structure of the BART, before and after which risk probabilities considerably differed. The change point occurred after one-sixth or one-third of trials in the first experiment, and between the initial and subsequent phases in the second experiment. Results of both experiments suggested that participants in all conditions adjusted their risk-taking behavior to the changed payoff structure of the task.

The results of the first experiment show that if initial experience indicated that high risk could be taken, risk-taking behavior decreased after the change in structure. If initial experience indicated that more cautious behavior was advantageous, risk-taking behavior increased after the change in structure. Importantly, the task structure after the change point was unbiased and comparable across experimental conditions, only the relative valence of events (balloon increase, balloon burst) differed as a function of previous experience. The results of the second experiment further highlighted that negative experience more strongly influenced the adjustment of behavior than positive experience: Slower adjustment was observed after negative events. Behavior adjustment after positive events did not reliably
differ from that of after baseline events. Nevertheless, due to participants’ updating of risk estimates in all conditions, their risk-taking behavior became similar by the end of the decision-making task.

**Interpretation of findings**

Regarding the analyses investigating the manipulated phases, we could conclude that initial manipulation was effective and occurred as planned in both experiments. Although manipulation took place in a separate phase in the second experiment, its effect seemed to be carried over. Thus, with this design, we tested the effect of initial experience in a controlled manner. If initial experience was not controlled, it could still influence findings of several BART studies, as illustrated by one of our previous experiments described in the Supplementary Material. According to the verbal reports, we found strategy-related differences only in the first experiment, where participants in the lucky long condition more often noticed that there was a switch in luck from early trials to later ones. Even in this condition, awareness was not associated with different risk-taking behavior. Therefore, the experimental design allowed us to test the *implicit* effect of initial experience and how participants change their behavior due to further experience.

It is a common approach when analyzing BART data to track how participants increase the number of balloon pumps in order to maximize reward after they have gained some experience with the task during earlier balloon trials. Therefore, we analyzed the main dependent variables of the task in five-balloon-long time bins. This also allowed us test whether the recent outcomes instead of the earlier ones would have had a more influential role in guiding behavior. Although both the MAP and the percentage of balloon bursts changed across the time bins, these changes followed different trajectory. In addition, their trajectories also differed between the two experiments, as detailed below.
In terms of the MAP, participants did not further enhance pumping behavior in any condition of the first experiment as the task progressed. Meanwhile, in the second experiment, pumping behavior strikingly increased in all conditions after experience with one-third of the task. This also meant that while lucky experience led to more cautious behavior after the end of manipulation in the first experiment, the opposite was observed in the second experiment. These between-experiments differences could be explained by the different balloon tolerances in the remaining and subsequent phases of each experiment. Particularly, in the second experiment, mean balloon tolerances were relatively high in Bin3, Bin4, and Bin5 (Fig. 3). Accordingly, balloon pumps were increased, and the frequency of balloon bursts was decreased in these time bins. Meanwhile, in the first experiment, mean balloon tolerances in the middle of the task were lower than in the second experiment; and, more importantly, these values were lower than over the initial balloons of the lucky long condition (Fig. 1). It is probable that participants reacted in a more sensitive manner to changes in the underlying structure when the manipulated phase was part of the main task; therefore, they did not pump the balloon to a larger size as the task progressed in the first experiment.

In terms of the changes in balloon bursts, it seems that participants’ behavior was predominantly determined by the most recent insights on balloon tolerances. As Figure 1 illustrates, when balloons tolerated a higher number of pumps in a given time bin, the frequency of balloon bursts was higher in the next time bin, even though the actual mean tolerance of this bin was lower. Then, risk taking was calibrated to the recent outcomes yielding a reduced frequency of balloon bursts in the next bin. This pattern was pronounced in the lucky long and lucky short conditions; and, it was also observable in the unlucky long and unlucky short conditions after participants had overcome initial experience. This bin-by-bin adaptation was not as emphasized in the second experiment as in the first one, possibly due to the above-described differences in balloon tolerances.
Although the MAP and the frequency of balloon bursts are related measures, the MAP might be less confounded by the balloon tolerances, since it more likely reflects the point until that a participant is willing to take risk. Considering this matter, to answer how long would it take to update risk estimates after initial experience, we base our interpretation on the changes of the MAP. In the first experiment, after manipulation had ended and participants had completed five more trials, behavior adjustment occurred. This was slightly different in the second experiment, where it took about ten balloons to start to adjust to the task and increase pumping. At the same time, dividing the task into five-balloon-long bins is arbitrary; therefore, based on these data, we could not directly determine the exact change point of behavior but could provide some approximation. It should also be noted that an exact change point might be less likely, instead, gradual behavior change might be assumed.

A related question is whether the length of manipulation influences the update of risk estimates. Indeed, we found overall higher MAP with short than long manipulation (significant main effect of Luck). However, our conservative conclusion would be that the different lengths of positive or negative experience would not considerably alter the updating process per se. In terms of the balloon bursts, we observed only a delay in behavior adjustment after long manipulation (i.e., the interval of the manipulation per se). Meanwhile, the MAP might have decreased in a milder manner after the lucky short than after the lucky long manipulation, maybe due to the smoother transition from the initial to the remainder phase. In particular, the short manipulation could have resulted in weaker memory traces on the initial payoff structure than the long one, which could have enabled the maintenance of a more risk-seeking behavior (cf. 10). These factors might explain that the lucky vs. unlucky difference persisted throughout the task in the case of short manipulation.

Behavior adjustment in the case of changing risks
The results of the first experiment are mostly in line with the findings of Koscielniak, et al.\textsuperscript{23}, regarding how the MAP changed throughout the task. However, while the between-conditions differences considerably decreased during the second half of our task in the case of long manipulation, these differences remained pronounced in the case of short manipulation. In line with the latter, the larger MAP after lucky than unlucky first trials also remained highly significant during the ten-balloon-long phases following manipulation in the study of Koscielniak, et al.\textsuperscript{23}. Our second experiment contradicts these findings as an increasing MAP was observed in the subsequent phase in all conditions, irrespective of valence.

As in our first experiment, the MAP was also overall reduced after prior losses as compared with standard experience in the study of Bonini, et al.\textsuperscript{26}. However, as the task progressed, participants of that study increased risk taking to recover from losses; and, this increase was steeper in the task version with prior losses. Similarly, our second experiment revealed that unlucky participants increased pumping as compared to their own level at the beginning of the task, probably to compensate for earlier losses, as explained by the prospect theory\textsuperscript{38,39}. However, the adjustment was slower in their case than that of the lucky participants, which is in line with the notion of loss aversion or the asymmetric sensitivity to gains and losses (cf.\textsuperscript{40}). Yet, these studies cannot be unequivocally compared as both Bonini, et al.\textsuperscript{26} and Koscielniak, et al.\textsuperscript{23} used within-subjects designs and different time bins that could have contributed to the persistence of the lucky/standard vs. unlucky difference.

Beyond the BART, other studies using various decision-making paradigms showed the profound influence of initial experience coupled with subjective feelings of luckiness on later risk taking\textsuperscript{41-44}. Similarly, adaptation to subsequent changes in dynamic environments was also influenced by the pervasive and often disadvantageous effect of prior events\textsuperscript{10-12}. Although we also emphasize the robust effect of initial experience, our results simultaneously suggest that participants continuously adjust their behavior to the most recent outcomes.
The simultaneous impact of recent outcomes could be reconciled with findings suggesting that forgetting the summary history of experience helps further adaptation\textsuperscript{10,11}. Particularly, in our experiments, participants saw only the accumulated total score and the score collected from the previous balloon on the screen; therefore, no detailed statistics were available about balloon tolerances. Although 60\% of strategic responses indicated a systematic approach, it is still conceivable that participants did not intensively track and memorize past events; instead, they changed their strategy in a dynamic manner, according to experiences on the latest balloons. This notion is supported by the observations of those studies that showed quick adaptation to changing underlying probabilities due to the continuous tracking of the recent outcomes\textsuperscript{45-47}. In addition, the study of Ashby and Rakow\textsuperscript{48} also indicated the prominent impact of recently observed outcomes for the sequential sampling of gambles and their valuation, modulated by individual differences in memory span (see also\textsuperscript{49}). Overall, the present findings together with the previous ones suggest that initial experience could be overcome in time by implicitly or explicitly sampling the underlying structure of the environmental stimuli\textsuperscript{9}.

**Limitations**

Although our findings clearly indicated the continuous adjustment of risk estimates in all conditions, the design of the experiments inherently limited what participants could initially learn about the balloons. Participants’ baseline belief about the pump number that balloons could have tolerated was not fixed a priori before the manipulation. They might have established this belief based on the experiences gathered during the first few trials. However, it was limited to what size they could pump the balloons at the beginning of the task; and, it was also unknown to what degree they were prone to pump these balloons (i.e., risk preferences). Relatedly, in contrast to previous studies on the effect of luck\textsuperscript{41-44}, we did not
investigate individual differences in beliefs about luck, and we did not ask in what degree participants felt themselves lucky after the entire task or after the initial phase. Similarly, we did not require participants to track and rate the confidence they had in their choice. We deliberately changed the underlying structure in an unsignaled and unexpected manner, which did not allow explicit questions about the sense of luck and/or luck beliefs. Although we still observed reasonable changes in behavior in line with changes in the structure, the cognitive processes underlying this behavior adaptation should be clarified in further studies.  
Relatedly, using computational models that account for the effect of initial experience should also be considered: With such analyses, initial experience per se could be quantified and other parameters of the task could be compared across participants and groups/conditions while controlling for this effect.

Conclusions

This study confirmed that early experience in sequential decision making profoundly influenced subsequent risky choices. Moreover, the results also highlighted that individuals could adjust their choice behavior by updating the risk estimates related to the changing environment. Adjustment occurred either by shifting choice behavior opposite to the direction oriented by early experience or by increasing overall risk taking. In addition, due to the differently changing payoff structures, choice behavior either remained risk averse after initially experiencing negative events or changed similarly as the task progressed, irrespective of the valence of initial experience. The speed of behavior adjustment was slower after negative events that after positive or neutral ones, suggesting the more influential role of the former in shaping choice behavior. Considering the observed effects of the continuously changing risks, choice behavior seems to be volatile rather than persistent.
Data availability

The datasets analysed during the current study are available from the corresponding authors on request.

References


Koscielniaik, M., Rydzewska, K. & Sedek, G. Effects of Age and Initial Risk Perception on Balloon Analog Risk Task: The Mediating Role of Processing Speed


**Acknowledgements**

This research was supported by the National Brain Research Program (project 2017-1.2.1-NKP-2017-00002, PI: D. N.); the Hungarian Scientific Research Fund (OTKA FK 124412, PI: A. K., OTKA PD 124148, PI: K. J., OTKA K 128016, PI: D. N.); the IDEXLYON Fellowship of the University of Lyon as part of the Programme Investissements d'Avenir (ANR-16-IDEX-0005 to D. N.); the János Bolyai Research Scholarship of the Hungarian Academy of Sciences (to A. K. and K. J.); and the Postdoctoral Fellowship of the Hungarian Academy of Sciences (to A. K.).

**Author contributions**

A. K. designed the study, created the scripts running the experiments, analyzed data, and wrote the manuscript; Z. K. designed the study, analyzed data, contributed to the interpretation of the results, wrote the manuscript, and critically revised previous versions of the manuscript; Á. T. designed the study, supervised data acquisition, analyzed data, wrote the manuscript, and critically revised previous versions of the manuscript; N. É. designed the
study, analyzed data of the explorative study presented in the Supplementary Material, contributed to the interpretation of the results, and critically revised previous versions of the manuscript; K. J. designed the study and critically revised previous versions of the manuscript, E. T-F. analyzed data, rated the interview questions, edited Supplementary Table S1, and revised previous versions of the manuscript; V. Cs. revised previous versions of the manuscript and provided intellectual support; D. N. designed the study, critically revised previous versions of the manuscript, and provided intellectual and financial support. All authors read and approved the final version of the manuscript.

**Competing interests**

The authors declare no competing interests.