

Supplementary Material

1 FULL LIST OF SURVEY ITEMS

Table S1. Full list of survey items. * marks the catch item in place to evaluate if users are still paying attention.

Item No.	Control group	CFE group	Response options
1	What do you think: Which plants were relevant to increase the number of Shubs in your pack? Please select ALL that you think were relevant.	What do you think: Which plants were relevant to increase the number of Shubs in your pack? Please select ALL that you think were relevant.	5 checkboxes, together with icons of the available plants + option "I do not know."
2	What do you think: Which plants were not relevant to increase the number of Shubs in your pack? Please select ALL that you think were not relevant.	What do you think: Which plants were not relevant to increase the number of Shubs in your pack? Please select ALL that you think were not relevant.	
3	I understood the overview of my past choices.	I understood the feedback on what choice would have led to a better result.	5 point Likert-scale, checkboxes with options: Strongly disagree - disagree - neutral - agree - strongly agree + option "I prefer not to answer."
4	I needed support to understand the overview of my past choices.	I needed support to understand the feedback on what choice would have led to a better result.	
5	I found that the overview of my past choices helped me to increase the number of Shubs.	I found that the feedback on what choice would have led to a better result helped me to increase the number of Shubs.	
6	I was able to use the overview of my past choices to increase the number of Shubs.	I was able to use the feedback on what choice would have led to a better result to increase the number of Shubs.	
7*	Are you still paying attention? If so, please select 'I prefer not to answer' for this question.	Are you still paying attention? If so, please select 'I prefer not to answer' for this question.	
8	I found inconsistencies in the overview of my past choices.	I found inconsistencies in the feedback on what choice would have led to a better result.	
9	I think most people would learn to work with the overview of their past choices very quickly.	I think most people would learn to work with the feedback on what choice would have led to a better result very quickly.	
10	I received the overview of my past choices in a timely and efficient manner.	I received the feedback on what choice would have led to a better result in a timely and efficient manner.	
Age	Please indicate your age.		Checkboxes with options: 18-24y, 25-34y, 35-44y, 45-54y, 55-64y, 65y and over
Gender	Which term most accurately describes your gender?		Checkboxes with options: Female, Male, Transgender female, Transgender male, Non-binary / gender non-conforming, Not listed, I prefer not to answer

2 EXEMPLARY USER JOURNEY THROUGH THE FIRST BLOCK OF THE GAME PHASE

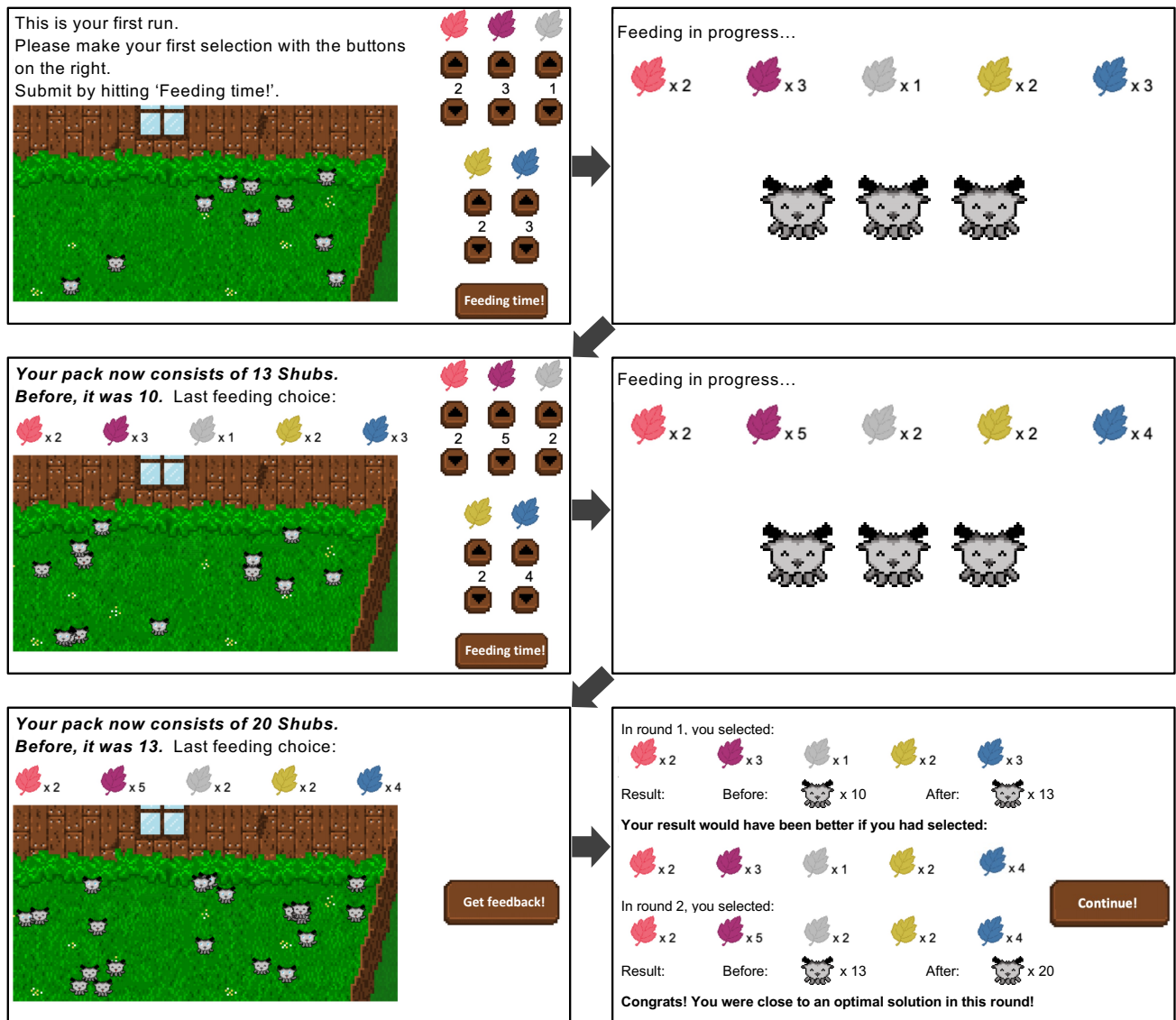


Figure S1. Exemplary user journey through the first block of the Alien Zoo game. Bold arrows indicate temporal succession of respective scenes. The figure highlights the iterative nature of the game with repeated user input and end-of-block presentation of counterfactual explanations (experimental group), or overview of past choices (control group). Note that plant counters are set to 0 at the beginning of each padlock scene. The figure displays the state after the exemplary user inserted their current choice. For this manuscript, font size in images of scenes was increased to improve visibility. Reprinted with permission from Kuhl, U., Artelt, A., and Hammer, B. (2022). Keep your friends close and your counterfactuals closer: Improved learning from closest rather than plausible counterfactual explanations in an abstract setting. In *2022 ACM Conference on Fairness, Accountability, and Transparency (FACCT '22)*, June 21–24, 2022, Seoul, Republic of Korea. doi:10.1145/3531146.353

3 DETAILED FEATURE RELEVANCE JUDGEMENTS

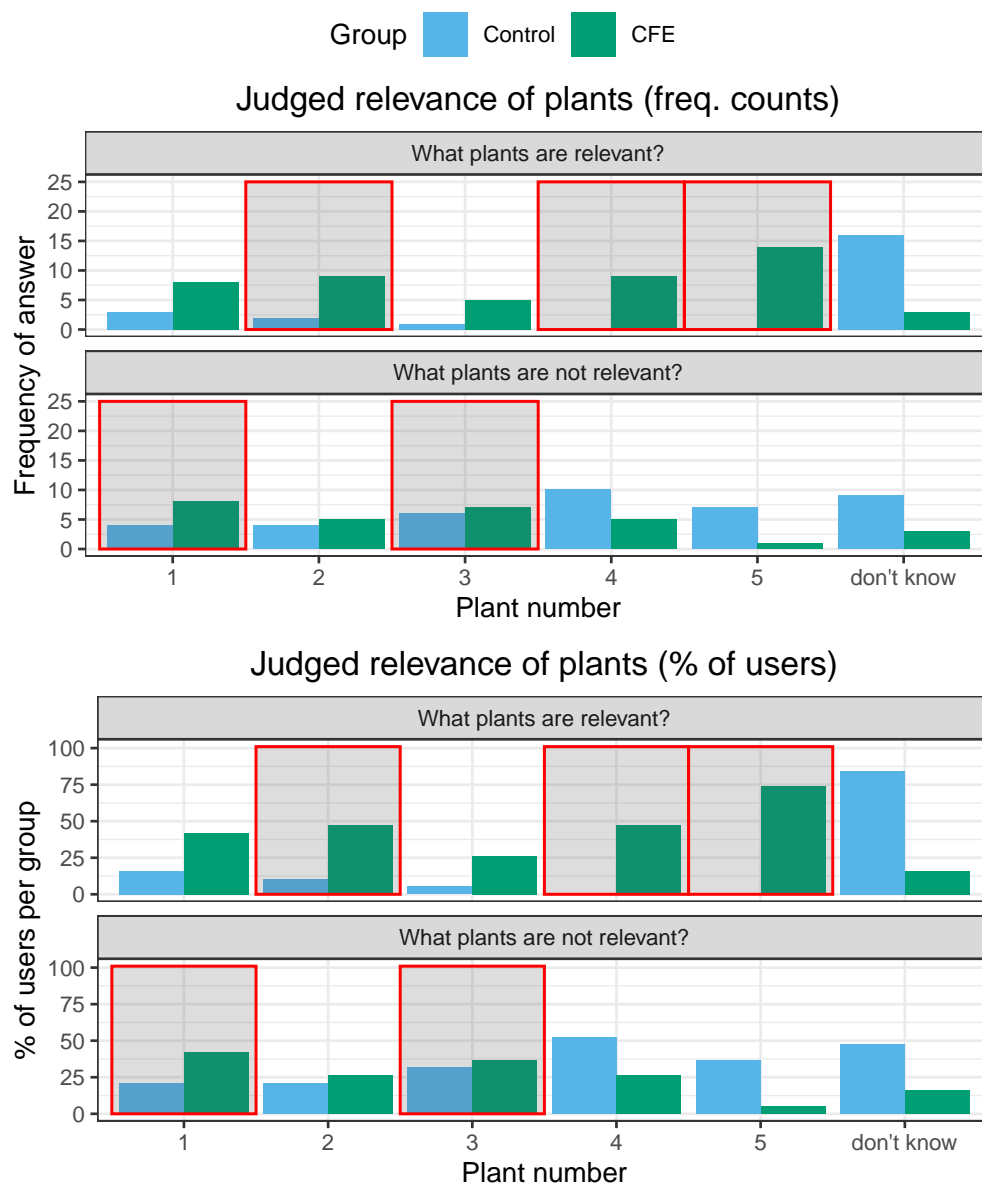


Figure S2. Experiment 1: Detailed relevance judgements per plant, as frequency of users (top), and percentage of users (bottom), respectively. Red boxes indicate relevant and non-relevant features.

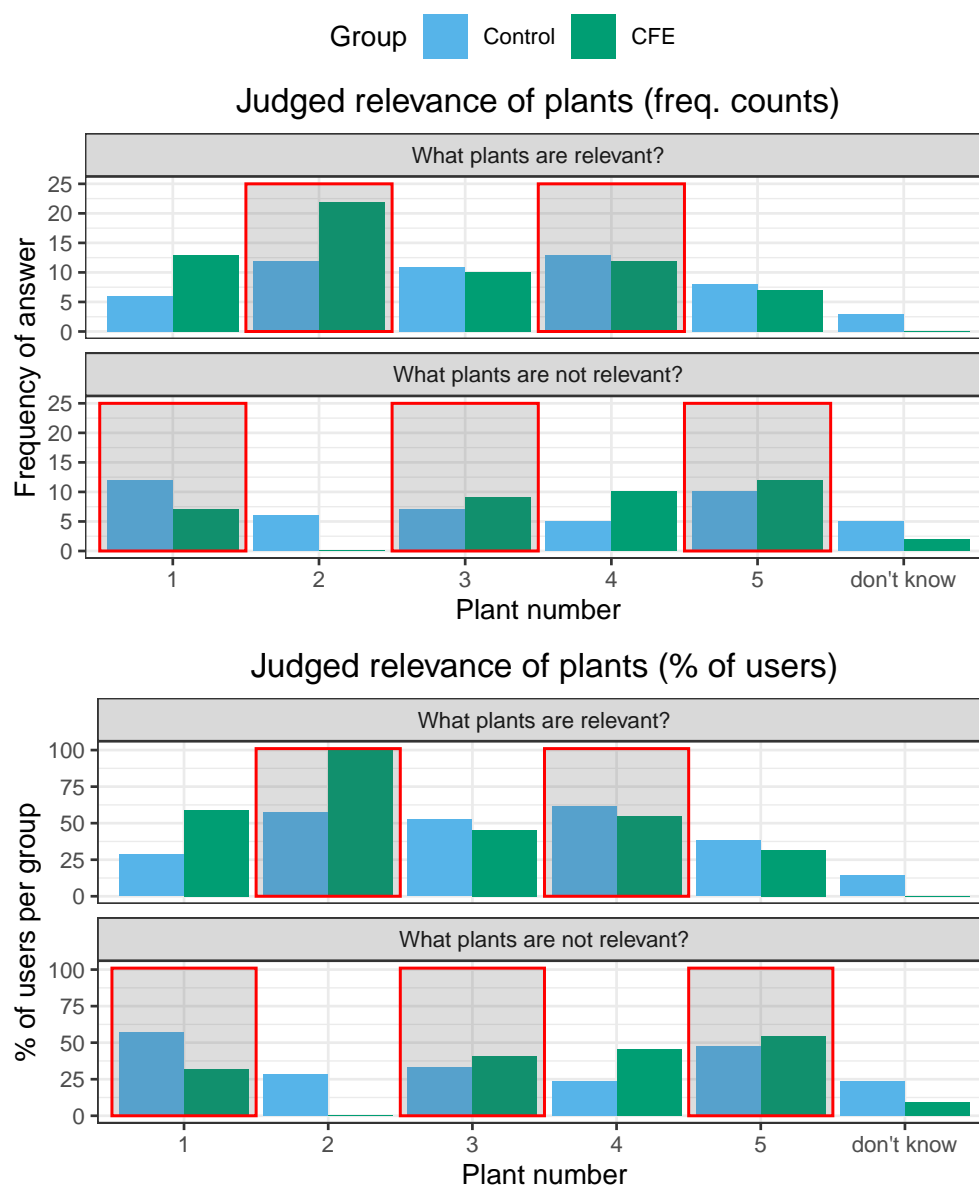


Figure S3. Experiment 2: Detailed relevance judgements per plant, as frequency of users (top), and percentage of users (bottom), respectively. Red boxes indicate relevant and non-relevant features.