The dissonance magnitude does not affect the way humans adapt to relational cognitive dissonances

Ron Hafzadi[1,*] **and Tomer Barak**[1]

[1]Professor Yonatan Loewenstein's lab, The Edmond and Lily Safra Center for Brain Sciences, The Hebrew University, Jerusalem, Israel
[*]ron.hafzadi@mail.huji.ac.il

ABSTRACT

Humans can resolve cognitive dissonances via several adaptation pathways. In relational cognitive dissonances, when an observation violates an expected relationship between objects, there are two types of adaptation pathways: adapting the expected relationship to match the observation or reinterpreting the observation to match the expected relationship. Previous research in Artificial Neural Networks (ANNs) has found that the dissonance magnitude is a crucial factor in determining the adaptation pathway in these dissonances: Large dissonances were resolved by reinterpreting the observation, and small dissonances were resolved by adjusting the expected relationship. To check whether this phenomenon exists in humans, we conducted an online experiment where the participants were induced with relational cognitive dissonances in several levels of magnitude. We found that the dissonance magnitude had no significant effect on the participants' adaptation pathways. In a second study, we examined whether performance anxiety prevented the participants from changing their initial choice, suppressing the reinterpretation adaptation pathway and covering the effect of the dissonance magnitude. We found that lower performance demands increased the selection of the reinterpretation adaptation pathway. However, this has not uncovered an effect of the dissonance magnitude. We discuss the inconsistency between our results and those obtained using ANNs.

Introduction

Cognitive dissonances, where attitudes, beliefs and behaviors come into conflict, can be resolved in a multitude of ways. The specific adaptation pathways that humans choose to resolve cognitive dissonances and their determinants are still not well understood[1].

An interesting case is the adaptation pathways of relational cognitive dissonances. In this case, dissonance arises due to an observation that violates an expected relationship. To illustrate this problem, consider a scientist who has consistently observed that particles of type B are larger than particles of type A. New experimental data, however, suggests the opposite: particles A are actually larger than particles B. These dissonances are of particular interest as there are two types of adaptation pathways that can resolve them: the scientist can either reinterpret the observation or modify the expected relationship. Reinterpreting the observation, the scientist can question the validity of the experimental results and maintain the view that B are larger than A. Alternatively, the scientist can modify the expected relationship, updating his view and conclude that particles of type A are indeed, larger than particles B.

Prior research on Artificial Neural Networks (ANNs) has found that the adaptation pathway chosen to resolve relational cognitive dissonance depends on the dissonance magnitude[2]. Specifically, they have constructed networks that learned a relationship between images. They then reversed that relationship, causing a relational dissonance, and evaluated how the networks adapted to this reversal. The ANNs had distinct relational and representational modules and could adapt either of these modules to resolve the dissonance. They found that the ANNs adapted their relational module (changing the expected relationship) when the dissonance was small and the representational module (reinterpreting the observation) when the dissonance was large.

To check whether the dissonance magnitude affects the adaptation pathway of humans when they resolve relational cognitive dissonances, we conducted an online study in which we exposed participants to such dissonances and controlled the dissonance magnitude. We found that unlike the ANNs, the dissonance magnitude had no effect on the adaptation resolution of humans. In the following sections, we present our experiments and results, and discuss possible explanations for the discrepancy between our results and those achieved by the ANNs.

Results

Color-ratio relation task

Participants were recruited to the study via Pavlovia (see Methods). We asked them to solve three tasks. Two tasks were used to evaluate their resolution pathways and we explain them in the next section. The primary task was the color-ratio *relation task*. This task's goal is to induce relational cognitive dissonances in the participants.

In the relation task (Fig. 1), the participants were shown two rectangles with blue and orange color filling. For brevity, we term these rectangles BORs (Blue-Orange Rectangles). The task was to select between the two BORs by clicking left or right using the keyboard. The two presented BORs had different ratios of blue and orange within them. The correct answer corresponded to the BOR with a higher ratio of one of the colors.

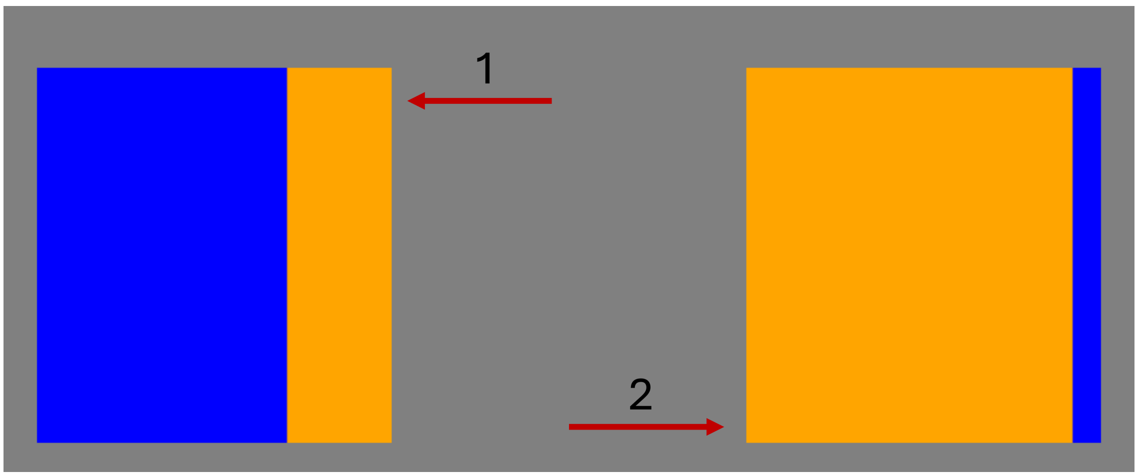

Figure 1. Relation task. Two BORs are presented. The difference between the BORs' blue color filling ratio is $\alpha = 0.625$. The correct answer depended on whether *more blue* or *more orange* was selected as the task rule. In the former case, the participants needed to click left (1), whereas in the latter case, right was the correct choice (2).

An important parameter in this task was the *ratio difference* between the correct and incorrect BORs, a parameter we termed α. Each participant was randomly assigned with a specific value of α for this task: the BORs that were presented to the participant were characterized by the same α, while the specific ratio values of the two BORs were chosen at random.

Initially, the participants did not know whether to select the rectangle with a higher ratio of blue (and a smaller ratio of orange) or a higher ratio of orange (and a smaller ratio of blue); they had to infer it by trial and error, based on feedback they received after each choice.

The color-ratio relation task continued until the participants reached 6 consecutive successes. We hypothesized that after reaching the success criteria, the participants had learned a relationship of the form "the BOR with *more/less blue/orange* is correct". To induce the participants with a relational cognitive dissonance, we reversed the rule that determined the correct answer. If BORs with more blue (or less orange) were the correct answers, now the BORs with less blue (or more orange) were correct.

The relationships the participants learned before the rule reversal consisted of two parts: a relational component (*more/less*) and the color focus (*blue/orange*). To resolve the dissonance following the rule reversal, the participants could either adapt the relational component (e.g., from more blue to less blue). Alternatively, they could maintain the relational component and change their color focus (e.g., from more blue to more orange). Based on the results obtained by the ANNs, we hypothesized that the selection between these two adaptation pathways would depend on the dissonance magnitude. In terms of the task parameter α, the larger α was, the larger was the difference between the rule before and after the dissonance. Therefore, we posited that α was a measure of the dissonance magnitude; we expected that small αs will lead to adaptation of the relationship, while big αs will lead to adaptation of the color focus.

Evaluating the adaptation pathways

To evaluate the adaptation pathway of the participants, we surrounded the relation task with two color-ratio *identification tasks*. The goal of the first identification task was to make the participants focus on a specific color when observing the BORs. Then, the participants performed the relation task, facing and resolving the cognitive dissonance. Finally, the last identification task was used to test whether the participants' color focus from the first identification task had changed.

In the first color-ratio identification task (Fig. 2), which participants solved before the relation task (Fig. 3a), the participants were shown a BOR, and their task was to click on a scale ranging from 0 to 1: the correct answer corresponded to the ratio of one of the colors. Initially, the participants did not know the color whose ratio determined the correct answer; they had to infer it by trial and error based on feedback after each item. We expected success in the identification task to result in the participants encoding the ratio of the relevant color whose ratio corresponded to the correct answer.

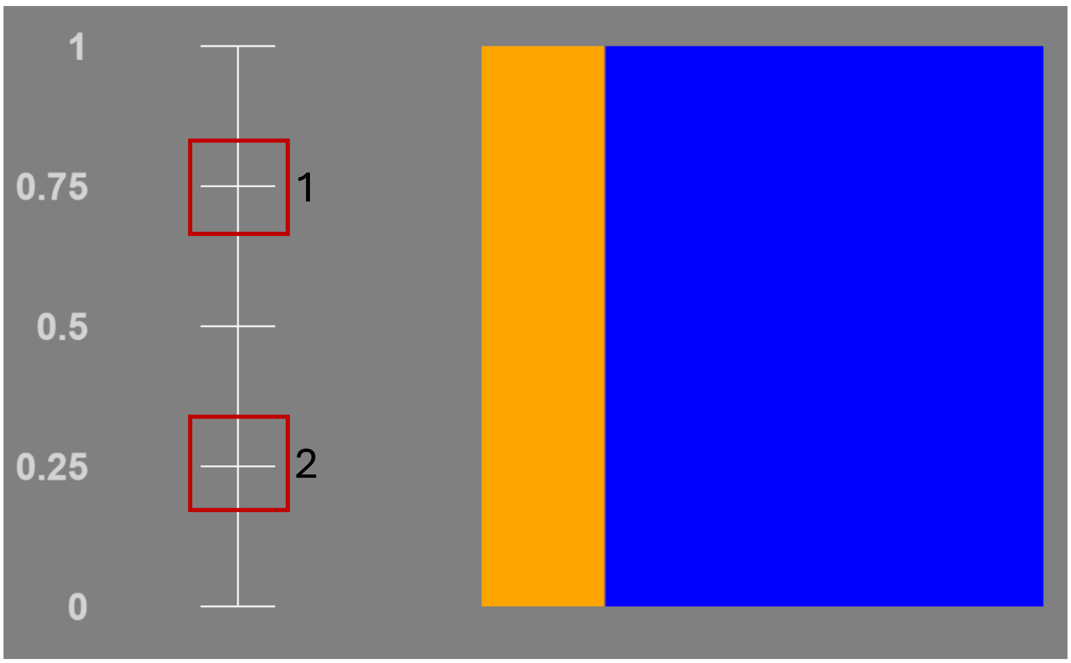

Figure 2. Identification task. A BOR was presented next to a slider. In this example, the BOR's blue ratio was 0.75. The participants were asked to click on the slider according to a rule that was related to the BOR. The correct slider clicking position (up to a tolerated error of 0.15) depended on whether the blue or orange ratio was selected as the task rule. In the former case, the participants needed to click within red box 1 on the slider, whereas in the latter case, red box 2 was the correct choice.

After completing the color-ratio identification task, the participants proceeded to the first phase of the color-ratio relation task (Fig. 3b). At this stage, we hypothesized that the participants' color focus in the first phase of the relation task would be the primary color focus they were trained on in the identification task. For example, a participant whose primary color focus from the identification task was blue would initially solve the relation task by selecting the BOR with a higher (or lower) ratio of blue.

Then, we inverted the rule, causing a relational cognitive dissonance for which the participants could adapt by either changing their color focus, or by adjusting the relationship (Fig. 3c). We assumed that if the color focus of the participants had changed, it would affect their response in the color-ratio identification task that followed.

In order to assess the color focus of the participants following the dissonance adaptation, we asked them to complete a single item of the identification task (Fig. 3d). In this final identification task, the blue ratio was 0.75. The participants did not know it was a single item, and that there was no correct or incorrect answer. We classified participants who clicked above 0.5 as blue-focused and below 0.5 as orange-focused. We checked the ratio of participants that switched their color focus between the first and final identification tasks. Our main hypothesis was that this ratio would correlate with the dissonance magnitude α.

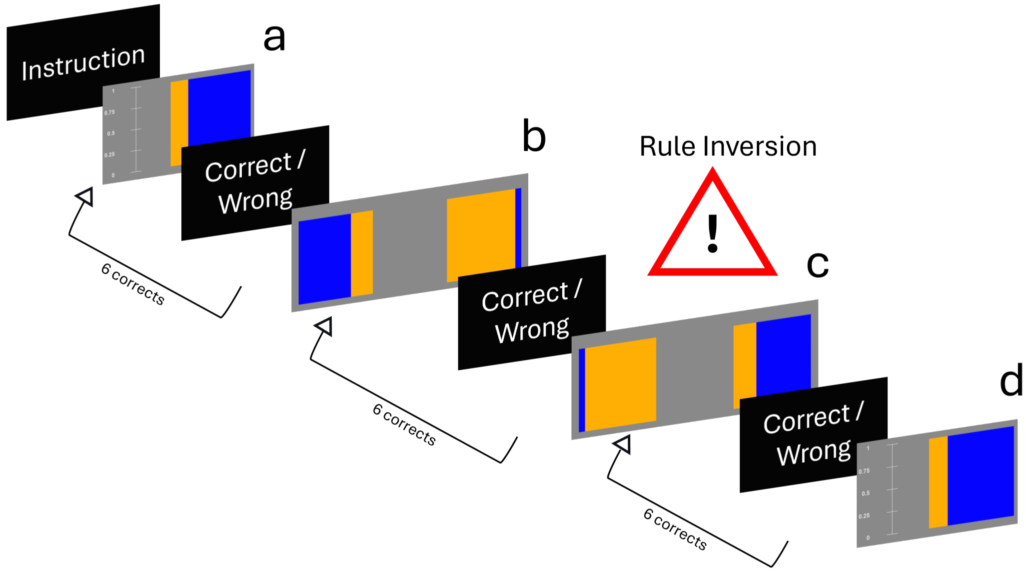

Figure 3. Experiment flow. **(a)** Color-ratio identification task. **(b)** Color-ratio relation task (1st phase). **(c)** Post inversion color-ratio relation task (2nd phase). **(d)** A single color-ratio identification task. The two identification tasks (**a** and **d**) were used to evaluate the participants' dissonance resolution pathway following the relational dissonance induced in the relation task (**b** and **c**).

Experiment 1 - The effect of the dissonance magnitude

61 participants solved the three tasks. Each participant was randomly assigned a specific value of α in the relation task, corresponding to a specific dissonance magnitude. We found that the ratio of participants that switched their color focus did not depend on the dissonance magnitude α (Fig. 4).

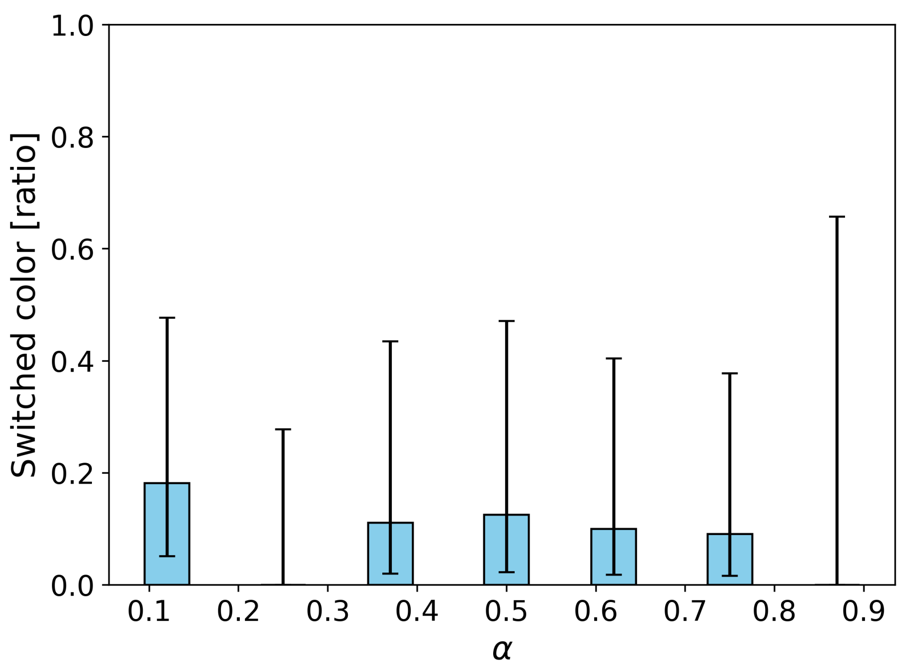

Figure 4. The dissonance magnitude effect on the adaptation pathway. The ratio of participants that changed their color focus between the first and last identification tasks did not depend on the dissonance magnitude α. Specifically, there was no significant difference between the switching color ratio of the small α group (0.12) and a large α group (0.75) ($t(20)=-0.6$; $p>0.05$). Error bars correspond to 95% CI (Wilson estimation).

Averaging over all conditions, most participants chose to adhere to the color they focused on in the first identification task (color switching ratio was 0.1 ± 0.1). We suspected that this conservative tendency was caused by performance anxiety. Specifically, in order to motivate the participants to finish the experiment, we promised a 1$ bonus for clicking on the "correct" answer in the final identification task. This bonus payment might have led the participants to be anxious when solving the final task, and stick to their previously working strategy. Therefore, to remove the effect of performance anxiety, we conducted another experiment without a bonus payment.

Experiment 2a - Without bonus payment

99 participants were assigned randomly to two groups of large ($\alpha = 0.87$) and small ($\alpha = 0.12$) dissonance magnitudes. In this experiment, there was no bonus promised for correctly solving the last identification task.

We expected that the lack of performance anxiety would lead to two related effects: the participants would more easily indicate that they changed their color focus (not risking the loss of money), and that it would uncover the effect of the dissonance magnitude. We found that our expectations were only partially met. The lack of bonus payment significantly increased the overall ratio of the participants who indicated that they had changed their primary color focus (Fig. 5a). However, this has not uncovered an effect of the dissonance magnitude. Even without the bonus payment, the dissonance magnitude did not affect the adaptation pathway of the participants (Fig. 5b).

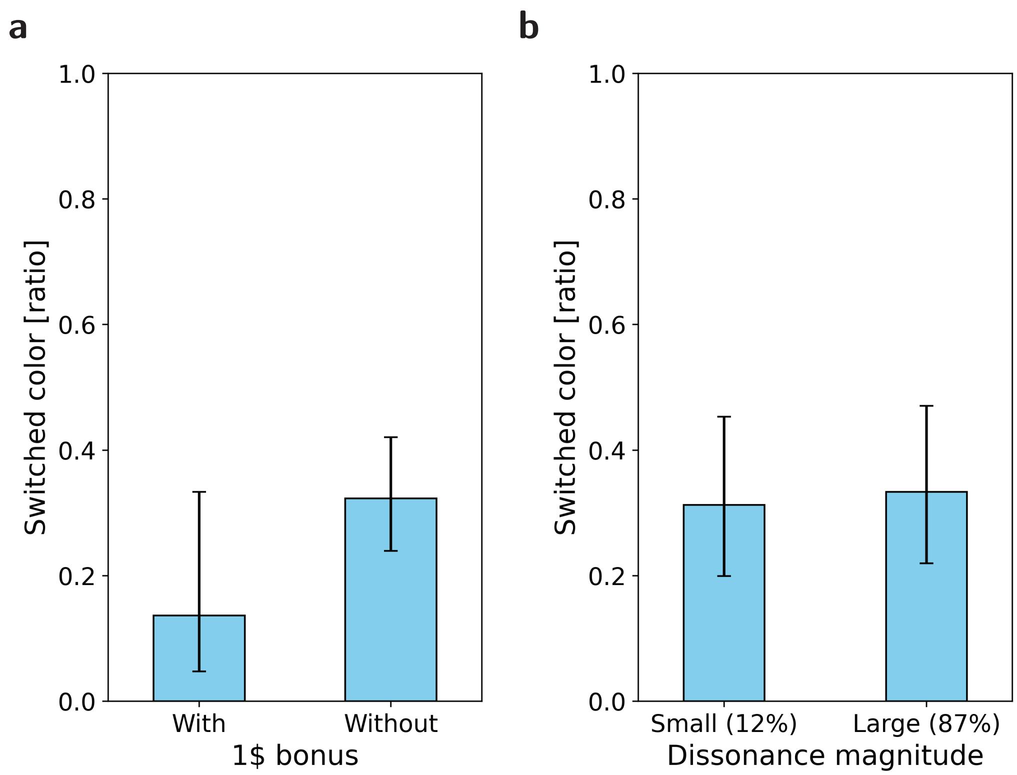

Figure 5. The bonus effect. (a) Removal of the monetary bonus led to a significantly higher ratio of participants' switching their color focus (t(119)=1.75; p=0.04). **(b)** Color switching was not affected by the dissonance magnitude in the bonus-less case (t(97)=0.2; p>0.05).

Experiment 2b - Implicit effects

Studies in the cognitive literature have shown that there are cases in which participants explicitly report one thing while implicit measures indicate that they believe another thing[3,4]. Since we saw that the participants' choice in the final identification task was affected by a monetary reward (5a), we suspected that their choice in this task reflected their explicit beliefs, while their implicit color focus might have been different.

Therefore, we checked for effects that the participants might have shown unknowingly which would indicate their implicit color focus. Specifically, in the final identification task, the participants have already learned that they can click on the slider with a rather large error margin (correct ratio ± 0.15). We hypothesized that the *exact* position in which they clicked on the slider might indicate their implicit beliefs. Specifically, we suspected that participants who did not explicitly switch their color focus, but did switch them implicitly, clicked closer to the non explicitly-chosen color. In terms of the dissonance magnitude, we expected that a large dissonance would cause more implicit color focus switches, and therefore more slider undershooting (clicking closer to the other color).

To test this hypothesis, we measured the extent of undershooting for participants who did not explicitly change their color focus. Consistent with our hypotheses, we found a significant ($t(65)=-2.55$; $p=0.006$) increase in undershooting when the dissonance magnitude was large (Fig. 6a). This result indicates that a large dissonance magnitude indeed leads to an increased rate of change in the color focus of the participants, and that it is an implicit effect. The cognitive literature suggests that a conflict between implicit and explicit beliefs is manifested by a larger processing time when solving a task[4]. In our case, *large* dissonance magnitudes changed the implicit, but not the explicit beliefs of the participants, increasing their implicit-explicit conflict. Consistent with the literature, we found that participants in the large dissonance magnitude condition responded slower in the final identification task, but the significance of this effect was not sufficient to reject the null hypothesis ($t(65)=1.5$; $p=0.07$) (Fig. 6b).

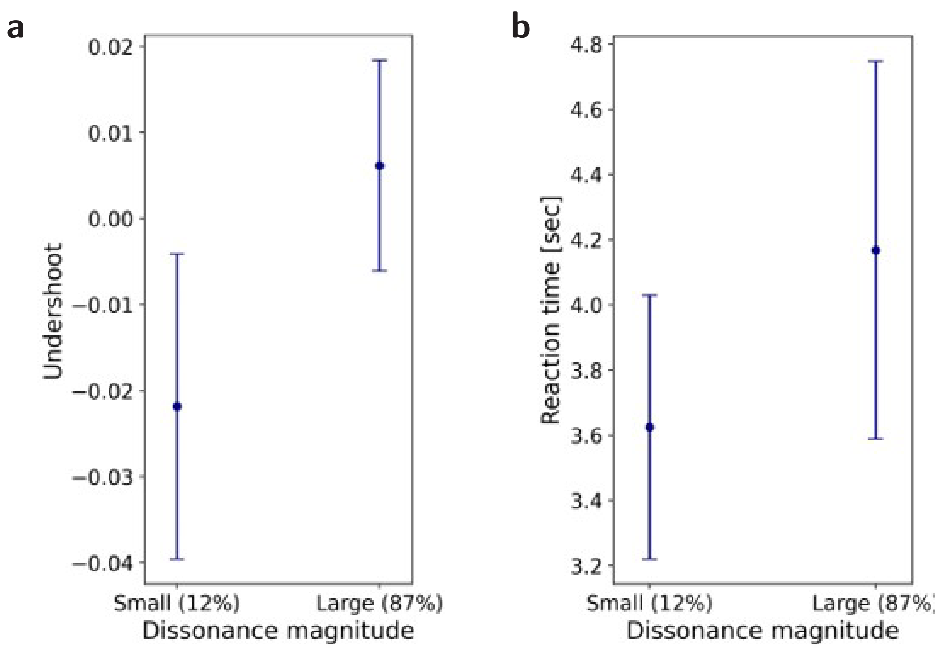

Figure 6. Implicit dissonance magnitude effects. (a) Slider undershooting in the final identification task for participants who did not switch their color focus. Undershoot measures by how much participants clicked closer to the color that was not their initial color focus. For example, if the initial color focus was blue, whose ratio in the final identification task is 0.75, undershoot corresponds to $0.75 - x$, where x is the participant's clicking position. **(b)** Reaction time in the final identification task for participants who did not switch their color focus.

Experiment 3 - Reconstruction

While the implicit measures indicated a dissonance magnitude effect, this effect was not what we initially hypothesized to find. Therefore, we suspected that our findings might have been a statistical error resulting from a multiple testing problem[5]. To validate our findings, we replicated the no-bonus experiment.

101 participants were assigned randomly to two groups of large ($\alpha = 0.87$) and small ($\alpha = 0.12$) dissonance magnitudes. We found that similar to the original study, the lack of a bonus payment has significantly ($p = 0.004$) increased the overall number of participants that switched their color focus compared to the participants from the first bonus-granting experiment (Fig. 7a). Also consistent with the initial study, the lack of bonus payment did not uncover an explicit dissonance magnitude effect on the color focus of the participants (Fig. 7b). However, unlike the initial study, we did not find the implicit undershooting and response time effects (Fig. 7c-d), indicating that these effects were indeed statistical errors.

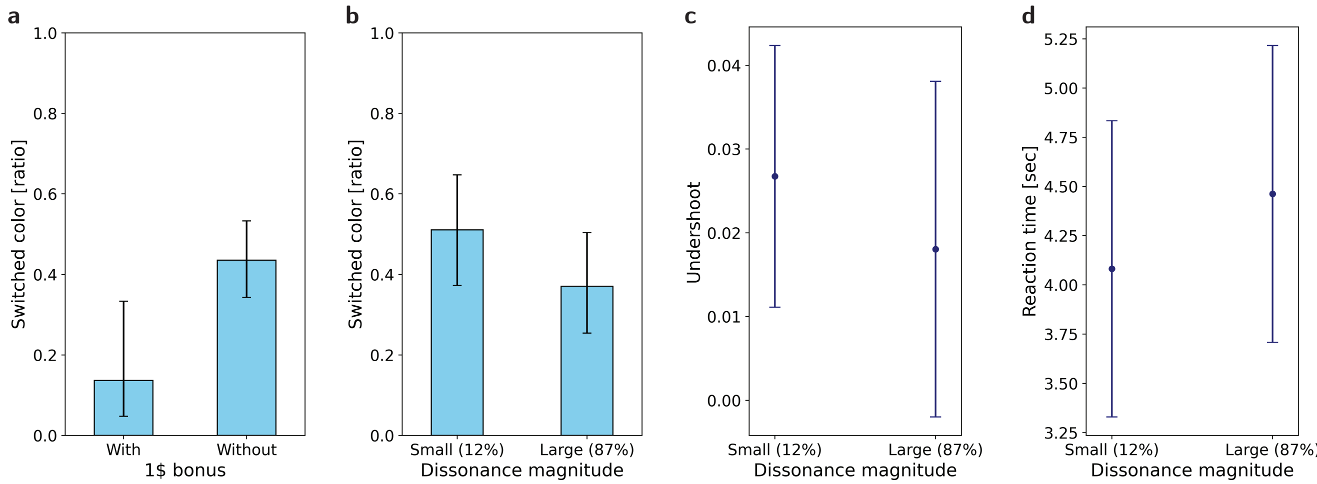

Figure 7. Experiment reconstruction. (a) The switched color ratio between the groups with and without the bonus. The results with bonus are taken from the first experiment and the results without bonus were obtained from the reconstruction experiment. The switched color ratio between the two groups is significantly different (t(121)=2.67; p=0.004). **(b)** The switched color ratio between the small and large dissonance groups. The two groups show no significant difference (t(99)=-1.42, p>0.05). **(c)** The Undershoot was not significantly different between the large and small dissonance groups (t(55)=0.62; p>0.05). **(d)** The reaction time was not significantly different between the large and small dissonance groups (t(55)=0.67, p>0.05).

Discussion

In our experiment, the dissonance magnitude had no effect on humans' resolution of relational cognitive dissonances. We built an online study in which participants were taught a specific relationship, and following their learning of this relationship, we inverted it. This created a relational cognitive dissonance whose magnitude corresponded to the relationship's (and its inverse) size. Relational cognitive dissonances can be resolved by readjusting the learned relationship or by reinterpreting the observation to match the learned relationship. Research on ANNs found that the dissonance magnitude was a key parameter for determining the resolution pathway, where a small dissonance magnitude led to an adaptation of the relationship and a large dissonance magnitude led to a reinterpretation of the observations. Unlike the ANNs, we found that the resolution pathways of human participants were independent of the dissonance magnitude.

Interestingly, the participants' preferred resolution pathway did depend on whether or not they received a bonus payment for successfully completing the experiment. This suggested that perhaps we measured an explicit decision of the participants, which is known to be affected by monetary reward, while the theory obtained for the ANNs holds for their implicit beliefs. Therefore, we measured implicit indications of the participants' resolution pathways. We initially found that the dissonance magnitude implicitly affected the participants' resolutions. However, this effect disappeared in a reconstruction of the experiment.

Between the two possible resolution pathways, we found that our participants tended to avoid reinterpreting the observation. These results can be explained by humans' risk aversion[6]. Reinterpreting the observation required participants to change the strategy that worked for them at the start of the experiment. Indeed, the tendency to stick with the working strategy was even more pronounced when a monetary reward was promised for a successful choice, as the risk was higher.

Overall, our experiment did not yield the dissonance magnitude effect predicted by the theory obtained for the ANNs. Like in the scientist example from the introduction, we can either explain this inconsistency by problems in the experiment or problems in the theory. As for our experiment, we relied on a few assumptions that may not have been entirely accurate. We assumed that the participants transferred their color focus from the first identification task to the relation task, and then transferred it again to the last identification task. However, the participants could have solved the tasks independently, focusing on a different color in each of the two task types (identification versus relation). This might explain the low ratio of changed color focus between the initial and final identification tasks and why this ratio did not depend on the relation task. An additional assumption was that we controlled the dissonance magnitude by changing the relational task's relationship size α. However, the participants might not have been sensitive to that relationship size. Our experimental settings allowed the participants to solve the relation task by only asking which BOR had more or less of a specific color, corresponding to the sign of α. They did not have to encode the specific difference size between the color ratios of the two BORs, corresponding to the size α. Not encoding the specific difference size is problematic from the perspective of the theory obtained for the ANNs; the networks' encoding of the size of α was crucial for the dissonance magnitude effect. If the participants had not encoded the size of α in our experiment, this explains why changing the size of α had no effect on their adaptation pathway.

In future research, for a better match with the theory, we suggest making sure that participants learn the specific relationship (by how much), and not a simplified relationship (greater than or less than). This can be achieved, for example, by a task in which learning the simplified relationship is not sufficient for succeeding in the task. Additionally, to verify the experimental control over the dissonance magnitude, we suggest testing well-known dissonance magnitude effects, such as changes in blood pressure or heart rate of the participants.

Another possibility for explaining the inconsistency between our results and the theory is that humans resolve cognitive dissonances by a different mechanism than the theory suggests. Specifically, unlike the ANNs used in the theory, humans are shaped not only by immediate input but also by extensive prior knowledge and experience, which can influence their preferred adaptation pathway. Additionally, human decision-making involves higher-order processes such as emotions (like risk aversion) and abstract reasoning, which extend beyond the capacities of the pattern recognition ANNs used in the theory. Thus, our study suggests that there is still a gap between theories derived from studying ANNs and human behavior. Future research could identify which part of the theory requires adjustment, advancing our understanding of how humans resolve the commonly experienced phenomenon of cognitive dissonance.

Methods

Participant details

261 participants were recruited to the study via Prolific. The participants were from the UK and USA, and between the ages of 18 and 60. They completed human recognition tests, and had a 100% reliability rating on the site. At the first experiment (n=61), participants were randomly assigned to a different α condition: 0.12; n=11, 0.25; n=10, 0.37; n=9, 0.5; n=8, 0.62; n=10, 0.75; n=11, 0.87; n=2. At the second experiment (n=99), participants were randomly assigned to either the small (.12; n=48) or large (.87; n=51) alpha condition. At the reconstruction experiment (n=101), participants were randomly assigned to either the small (.12; n=47) or large (.87; n=54) alpha condition. The Data was collected at the end of the fall semester in 2024.

Experiment details

The experiment was created using PsychoPy[7]. The experiment starts with asking the participants for consent to participate in the experiment. Once a participant had clicked that he is ready to participate, we start the experiment with the identification task. The identification task is constructed of two elements, a BOR, and a slider. All BORs were created using a simple python script, that creates a 200x200 pixels rectangles filled with blue and orange with random ratios. The slider was created using the built-in Psychopy slider tool. Using the ratios of blue-orange of the BORs, we programmed the slider to display a positive feedback (the word "correct") for clicking $\pm.15$ from the true ratio, and to display a negative feedback (the word "wrong") otherwise. This ratio was based on the primary color focus given to the participants, which was randomly chosen between blue and orange at the start of the experiment. This task was in a loop until the participants reached the success criteria – 6 consecutive correct answers. Once reached, the participants moved to the color-ratio relation task. The relation task is constructed of 2 BORs. We created a folder for each α, that holds a large number of α twin BORs (an image with X% blue, and an image with X + α % blue). The participants were randomly assigned to one of these folders at the start of the experiment, and got displayed with 2 twin BORs of the specific α folder. The participants had to click left or right using their keyboard. Using the Psychopy's built-in keyboard tool, we used their input to check whether it was pointed at the excess α BOR or the deficit α BOR. Once the participants had learned the assigned rule, and got to 6-consecutive correct answers, we inverted that rule. The rule inversion was made by flipping the programmed correct answer BOR: If pre-inversion the correct answer referred to the excess α BOR, now the correct answer was the deficit α BOR, and vice-versa. Once 6-consecutive correct answers were made, the participants moved to the final identification task. The final identification task comprised of a deterministic 0.75 blue BOR and a slider. If participants clicked above 0.5 we saved their answer as blue, otherwise as orange. Throughout this study, we used α values that ranged from 0.12 to 0.87. The left-right arrangement of blue and orange within each BOR was random. In the relation task, the correct answer BOR was presented either on the left or right side at random.

Instructions

At the start of the experiment "Welcome. In this experiment you are requested to identify rules that are associated with blue and orange rectangles presented on the screen. You will need to select an answer based on the ratios of orange and blue in the rectangles. After each choice, you will get a feedback. Try to identify the rules based on these feedbacks. Please follow the specific instructions of each part."

Before the first identification task "In this part, you will need to click on the slider with your mouse."

Before the relation task "In this part, you will need to choose left or right using the arrows of your keyboard."

In the bonus version of the experiment, before the final identification task "Now you are ready for the final and most important task. You will have only one chance to click on the slider in the correct position. If you get it right, you will get the bonus payment."

In the bonus-less version of the experiment, before the final identification task "Now you are ready for the final task. It is the slider task with a single item".

References

1. McGrath, A. Dealing with dissonance: A review of cognitive dissonance reduction. *Soc. Pers. Psychol. Compass* **11**, e12362, DOI: 10.1111/spc3.12362 (2017).
2. Barak, T. & Loewenstein, Y. Is it me, or is A larger than B: uncovering the determinants of relational cognitive dissonance resolution. *In preparation* (2024).
3. Greenwald, A. G., McGhee, D. E. & Schwartz, J. L. Measuring individual differences in implicit cognition: the implicit association test. *J. Pers. Soc. Psychol.* **74**, 1464–1480, DOI: 10.1037//0022-3514.74.6.1464 (1998).
4. Karpinski, A. & Hilton, J. L. Attitudes and the Implicit Association Test. *J. Pers. Soc. Psychol.* **81**, 774–788, DOI: 10.1037/0022-3514.81.5.774 (2001). Place: US Publisher: American Psychological Association.
5. Armstrong, R. A. When to use the Bonferroni correction. *Ophthalmic Physiol. Opt.* **34**, 502–508, DOI: 10.1111/opo.12131 (2014). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/opo.12131.
6. Holt, C. A. & Laury, S. K. Risk Aversion and Incentive Effects. *Am. Econ. Rev.* **92**, 1644–1655, DOI: 10.1257/000282802762024700 (2002).
7. Peirce, J. *et al.* PsychoPy2: Experiments in behavior made easy. *Behav. Res. Methods* **51**, 195–203, DOI: 10.3758/s13428-018-01193-y (2019).