

Dr. Chris Chambers
Recommender
PCI Registered Reports

January 2, 2023

Dear Dr. Chris Chambers,

Thank you for your time and effort. We appreciate the reviewers' thoughtful and constructive comments regarding our Stage 1 manuscript titled "The Medusa effect: A registered replication report of Will, Merritt, Jenkins, and Kingstone (2021)."

We are pleased with your prompt action and the reviewers' valuable comments. We have substantially revised our Stage 1 manuscript based on the reviewers' comments and our discussion. We believe we have addressed every point raised in the review comments, and thanks to this revision, we have modified our manuscript into a much-improved version. We have also uploaded a copy of the original manuscript with all the changes colored in blue.

Replies to the comments by Dr. Chris Chambers

I have now obtained four very helpful and constructive reviews of your Stage 1 submission. As you will see, the reviews are overall positive and, in my own reading, I found myself agreeing with their general enthusiasm for replicating this intriguing phenomenon. As is often the case with Registered Reports, the reviews highlight a number of areas that would benefit from clarification and possible design amendments, ensuring that the replication is as well motivated and diagnostic as possible about the replicability of the original study. In revising, foremost issues to consider are deviations from the original methodology (which should be minimised as much as possible, and clearly justified where required), strengthening the motivation for the replication (which should be straightforward to achieve), sufficiency of control conditions, addition of key details concerning the procedures and analysis plans, and clearly stating the conditions under which the results would be deemed to constitute a successful replication of the original findings.

On this basis I am happy to invite a thorough revision and response, which I will likely return to a subset of the reviewers for re-evaluation.

Reply:

We really appreciate your thoughtful and constructive comments. We are sorry for the careless mistakes and explanations that were not detailed enough. We have now discussed the comments and revised our manuscript based on all the thought-provoking suggestions—in particular, about the motivation for replication and the composition of our study. We hope that

the revised manuscript is easier to read and that the research presents an adequate replication design.

Replies to the peer review comments by Dr. Alan Kingstone

1-1

STUDY 1

1. For Study 1 (replication of Will et al. Experiment 2), as noted in the Will et al. (2021) supplementary material (see below as well), Will et al. included with each image and question a short definition of the mind perception conditions: Agency (ability to do) and Experience (ability to feel). It's not clear that the present Study 1 is including the short definitions with the images. It might be useful to include them as the longer definitions provided at the start of the study can be a little confusing to participants and hard for them to remember.

Short Definitions Included with the Questions of Agency and Experience

Agency condition

Which person seems to have more Agency (ability to do)?

Experience condition

Which person seems to have more Experience (ability to feel)?

Reply:

We appreciate your careful consideration. Since the procedure of our Studies 1 and 2 will be the same as the original research, short definitions of Agency and Experience conditions will be included with each image and question. We have added this detail into our manuscript as follows:

[p.13]

The question at the top will be about the person on the left side of the picture (e.g., Please rate the Experience (ability to feel) of the person on the left), and the one at the bottom will refer to the person on the right side of the screen (e.g., Please rate the Experience (ability to feel) of the person on the right).

[p.17]

Participants will complete a two-alternative forced choice (2AFC) task by answering three questions: Which person seems more real? Which person seems to have more Agency (ability to do)? Which person seems to have more Experience (ability to feel)?

1-2

STUDY 2: Dictator Game

In order to look at individual differences, Will et al. (2021, Experiment 5) had participants first do a two-alternative forced-choice (2AFC) task, identifying which of the two people in the photo seem more Real and have more Agency and Experience. And only then,

afterwards, did participants complete a one-shot dictator game task. It is not clear that this methodology will be followed in Study 2, for the following reasons.

A. In the current Study 2 design, the authors *do* mention pictorial abstraction (Realness, Agency, or Experience), but they state that it will be a between-subject factor, not a within subject factor as was the case in Will et al. Experiment 5.

B. The authors make no mention of how Realness, Agency, and Experience will be assessed (as noted, in Will et al, Experiment 5, it was with a 2AFC task).

C. The procedure and data analysis sections make no mention of any measurement of pictorial abstraction at all, and just discuss the Dictator Game.

If my current reading of Study 2 is correct, then I would encourage the authors to consider doing the within-subject 2AFC task before the Dictator game, as it is then a full replication of Will et al.'s Experiment 5. It also carries with it the additional benefit of allowing one to measure individual differences. Finally, we have recently collected data that suggests the Dictator Game effect may be larger when participants first assess the images for realness, agency, and experience.

Reply:

Thank you so much for pointing this out and providing the information.

First, we would like to apologize for the critical careless mistake. The pictorial abstraction is indeed a within-subject factor rather than a between-subject factor in both Study 1 and Study 2. Since it is a direct replication, the design and procedure of our studies will be the same as the original research. We have corrected this mistake in our revised manuscript.

As for omitting the 2AFC task before the dictator game in Study 2, we have discussed it again based on your careful consideration and information and decided to add it to our revised manuscript. Our initial consideration was that the purpose of the 2AFC task of the original Experiment was only to confirm the differentiation of mind perception between L1 and L2 of the stimuli. Considering a mind perception task before the dictator game is a requisite for relating behavior to perception, we have accepted your suggestion and conduct the within-subject 2AFC task in the first step of our Study 2. As a result, the procedure and method will be the same as Will et al.'s (2021) Experiment 5.

We would like to thank you once again for providing us such important information, and we have revised our manuscript according to your suggestions.

1-3

Minor comment

Page 6 "In conclusion, the original study The dictator game principle was upheld throughout the process, namely, that the recipient did not influence the size of the allocation, and since the participants' actions were fully autonomous, they could reflect their own wishes in the experiment."

- I didn't quite understand what was meant by the underlined section. For example, in Will et al. (2021) the recipient (L1 vs L2) did influence the size of the allocation.

Reply:

Thank you for pointing this out. We apologize for the confusing content. This sentence meant to express that since the recipients in the dictator game are fictitious, the participants would not feel any ethical pressure, and their actions would be autonomous. It may further illustrate that the recipients' characteristics could be explained as the results of the dictator game. We have modified this in our revised manuscript.

Replies to the peer review comments by Dr. Brittany Cassidy

2-1

I do wonder if the authors should go ahead and complete the conditional experiments—it would be meaningful to know if the Medusa effect replicates across diverse image sets cross-culturally.

Reply:

Thank you so much for your consideration. We are afraid that we will only conduct Studies 1 and 2 in the first phase, since the very first purpose of the present research is to examine whether the Medusa effect itself could be replicated in Japan by using exactly the same stimuli as the original experiments to explore the cross-culture scope of the abstraction costs across different levels. If the replication succeeded, not only the existence but also the scope of the effect would be confirmed again by using the cross-cultural stimuli. If the replication did not succeed, we would conduct the conditional study, which refers to our Study 3, to further discuss the difference and possible causes of these various results.

2-2

Overall, I am enthusiastic about pre-registered replications. I do think, however, that the authors could be stronger about how their proposed work replicates *and* extends the literature. For example, the authors could be much stronger about how it is meaningful for the literature for findings to be replicated across diverse cultures. They touch on it some, but leave their thoughts a bit vague and just describe “generalizability.” But, they are selling themselves short, I think, by downplaying the potential for the Medusa effect as culturally generalizable. This may be important to bring up considering much work showing perceptual and social cognitive cross-cultural differences. Generalizability would be especially cool here.

Reply:

Thank you very much for your understanding of the purpose of our research and your thought-provoking comments. These suggestions are particularly helpful for revising and improving our paper. We have modified our manuscript by emphasizing the importance of our research and giving more detailed statements of how meaningful it is according to your suggestions.

As you have mentioned in the comment, the exploration of cross-culture in the Medusa effect is critical for our replication research. Since the Medusa effect refers to a difference in the mind perception of others during the process of pictorial abstraction, the factors that could affect

social cognition should be playing a key role in it. We have modified our revised manuscript as follows:

[p.8]

Furthermore, whether this effect can be replicated in a different group of participants is also important for its generalization. The original dictator game data suggest that the susceptibility to the Medusa effect may vary among different people (Will et al., 2021). This result may reflect individual differences in underlying cognitive abilities, such as face perception ability (Frischen et al., 2007). Moreover, when it comes to the cross-culture effect on mind perception, according to Krumhuber et al. (2015), when participants evaluate mind perception of faces that range on a continuum from real to artificial, intergroup processes (i.e., in-group favoritism and out-group dehumanization) play a key role in humans' perception. To be specific, for instance, participants from India evaluated South Asian (in-group) faces more mindful than Caucasian (out-group) faces (Krumhuber et al., 2015). This suggests that a similar influence may also exist when it comes to pictorial abstraction perception. Since the original stimuli are "out-group" faces for participants in Japan, based on the previous research, the cross-culture effect may be one of the reasons to account for less sensitivity of mind perception.

2-3

I also think the authors could be stronger in motivating why *this* particular effect should be replicated. Replication is a good thing altogether, but why *this* effect out of the myriad effects in the literature? What makes this effect especially relevant/important to pay extra attention to?

Reply:

We are grateful for this valuable suggestion. We have added a more detailed explanation for the reason why we decided to pay attention to the Medusa effect and conduct the replication according to your comments in the revised manuscript as follows:

[p.7]

The original research was the first to reveal the Medusa effect and explore its prosocial effects, showing that the abstraction levels of the picture itself can reduce mind perception. This finding of the abstraction cost between different levels of pictorial abstraction suggests a novel hypothesis for future researchers, namely that people's cognitive effects may be weakened by a higher level of pictorial abstraction; in this research, for example, mind perception is reduced. We noted that, since pictorial representation has been playing a vital role in providing information on the Internet (Latha & Fathima, 2019), the differences in abstraction levels could affect considerable social interactions by influencing mind perception; for instance, they may affect online experiments involving pictorial stimuli of faces.

Replies to the peer review comments by anonymous Reviewer 1

3-1

My main comment is about the possibility of not observing the ‘Medusa’ effect in Exps. 1 and 2. The authors identified two main possibilities: 1) the Medusa effect does not exist, or it exists only under very limited conditions (pages 15-16). 2) The original stimuli were not adequate to detect the effect. In both cases, I would ask the authors to provide clearer explanations; in particular, for 1) please clarify what you mean by ‘very limited conditions’. As for 2) I do agree that ethnic membership of original stimuli (White) and the participants that will be recruited in Exps. 3a/3b could be confounding, as we know that social perception is deeply shaped by the ethnicity of both the stimulus and the observer. Nevertheless, also for 2) please clarify what you mean, by providing a stronger rationale supported by references.

In any case, I also wonder if the (potential) absence of the ‘Medusa effect’ within Japanese people may be due to cultural differences that are known to exist in Western and Eastern cultures, such as differences in perception of social scenes (see, e.g., Masuda, 2017) and more in general differences in attitudes towards people (Western and Eastern societies are more individualistic and collectivist, respectively). For instance, while Westerners can focus on specific stimuli, Easterners are more influenced by the presence of non-target social stimuli presented in peripheral positions (and therefore this could influence the rating associated with L1 or L2 levels? This is just a speculative interpretation, but I would be happy to hear some comments).

Reply:

Thank you for pointing this out.

1) Regarding the “very limited conditions,” we mean that except for pictorial abstraction levels, there may also be some possible factors we did not discuss in this research on the Medusa effect. For example, mind perception may vary among different participants on account of their preferences for facial appearance.

2) For the possible ethnic limitations of the original stimuli, in addition to the well-known cultural differences between Western and Eastern cultures in the perception of social scenes you have mentioned in the comment, previous research has also found intergroup biases in mind perception. We have added further explanations in our revised manuscript as follows:

[p.8]

Moreover, when it comes to the cross-culture effect on mind perception, according to Krumhuber et al. (2015), when participants evaluate mind perception of faces that range on a continuum from real to artificial, intergroup processes (i.e., in-group favoritism and out-group dehumanization) play a key role in humans’ perception. To be specific, for instance, participants from India evaluated South Asian (in-group) faces more mindful than Caucasian (out-group) faces (Krumhuber et al., 2015). This suggests that a similar influence may also exist when it comes to pictorial abstraction perception. Since the original stimuli are “out-group” faces for participants in Japan, based on the previous research, the cross-culture effect may be one of the reasons to account for less sensitivity of mind perception.

3-2

My second comment is about the sample: You declare that people 18-99 years old will be included, but I am wondering if an upper limit would be reasonable, to get more homogeneous samples and avoid possible differences in social perception that may arise across people of different ages.

Reply:

Thank you for pointing this out. Regarding the upper age limit we set in the manuscript, the intention was to not to restrict the study to a specific age group, since the original research did not restrict participants' age. However, as we will recruit participants through an online crowdsourcing platform, the upper age refers to a recruitment principle by which the participants should have no cognitive impairment. Nevertheless, considering that this may be misleading, we have decided to delete it from our revised manuscript.

Replies to the peer review comments by anonymous Reviewer 2

4-1

The motivation for conducting the present study is simply that the Medusa effect has never been replicated before (p. 6). While I am in perfect agreement with the authors that the Medusa effect is interesting and important, this rationale of course applies to many (if not most) effects that are just as interesting and important. And actually, the original study already contains multiple direct and indirect demonstrations of the effect, and follows open data practices. More generally, this is not a direct replication testing the reproducibility of the Medusa effect, because it crucially involves a different sample/culture. This should be highlighted and motivated as a key difference: is there past work suggesting cultural differences in mind perception or abstraction perception? Do the authors have reason to believe this will make no difference?

Reply:

Thank you very much for your understanding of the purpose of our research and your thoughtful comments. We apologize for not giving a detailed enough explanation of our motivation for replicating the original research in our manuscript.

As you have mentioned in the comment, the original study conducted five experiments to demonstrate that the effect (picture abstraction cost) exists and its prosocial behavior effect also exists, which made this effect quite intriguing and reliable from our point of view. Furthermore, when it comes to the cross-culture effect on mind perception, according to Krumhuber et al. (2015), when participants evaluate the mind perception of faces that range on a continuum from real to artificial, intergroup processes (i.e., in-group favoritism and out-group dehumanization) play a key role in the perception of humans. To be specific, for instance, participants from India evaluated South Asian (in-group) faces as more mindful than Caucasian (out-group) faces. This finding suggests that a similar influence may also exist when it comes to pictorial abstraction perception. Since the original stimuli are "out-group" faces for

participants in Japan, based on the previous research, the cross-culture effect may be one of the reasons to account for less sensitivity of mind perception.

We have modified this in our revised manuscript and further explained our motivation for conducting the present study.

4-2

I had several comments on the discussion of possible outcomes:

Will the authors take their results to support H1 and H2 only if all three DVs show an effect? What if only two do? Or one?

Reply:

Thank you for your comments. This comment would refer to the principle of supporting H1 of our Study 1 and H2-a of our Study 2. According to the original research, Realness, Agency, and Experience are three dimensions of the quantitative framework to measure our DV, which is mind perception. We will consider that the results support H1 and H2-a only if all these three show significant differences ($\alpha = .05$). If only one or two dimensions show an effect, we will consider that the results do not support H1 and H2-a, and the differentiation of mind perception between L1 and L2 cannot be confirmed.

In this regard, we have made it clear in our revised manuscript as follows:

[p.8]

For H1 (in Study 1), participants would rate L1 as having higher levels of both Realness, Agency, and Experience than L2.

Regarding H2-a, participants would choose L1 rather than L2 in all three Realness, Agency, and Experience dimensions.

4-3

The authors only mention the possibility of failure to replicate the original effect ("If neither H1 nor H2 are supported, the reproducibility of the Medusa effect may be problematic.") but do not discuss it further. This might instead be a good occasion to mention differences between the original and current study (e.g., different sample, different recruitment platform, ...).

Reply:

Thank you for your suggestion. We are excited to accept your suggestion, and we have further discussed the differences between the original study and ours that may lead to different results in our revised manuscript as follows:

[p.18]

Study 3 is a conditional study comprising two parts, Studies 3-a and 3-b, and will be conducted only if H1 and/or H2 was not supported in Study 1 and/or Study 2. There may be two possible reasons for this failure. First, there is a possibility that the Medusa effect does not exist or only exists under very limited conditions; for example, the results may vary among different participants on account of their preferences for facial appearance. Secondly, the results may

have been caused by limitations of the stimuli (e.g., quality and race). Considering the above reasons, we planned this conditional study (Study 3) by using newly designed stimuli with higher quality and the same ethnicity of participants to further replicate the original Experiments 2 and 5.

4-4

It is unclear why limitations of the stimuli may explain disconfirmation of H1 and confirmation of H2. Why would the stimuli matter for H1, but not H2? And why would they explain a disconfirmation, if they are the same as in the original study?

Reply:

Thank you for pointing this out. There are two main differences in confirming H1 and H2-a (the mind perception measurement). First, Study 1 uses a rating task, but Study 2-a uses a two-alternative forced choice (2AFC) task. Since the quantitative frameworks are the same (refer to three dimensions of Experience, Agency, and Realness), we do not think this would be the potential reason for the possible different results. Secondly, the stimuli are different. In Study 1, stimuli are pictures downloaded from the Internet, which vary in terms of size, quality, gender, race, and emotional expression. In contrast, the stimuli of Study 2 are newly made presentations with better control of irrelevant variables, including equated facial appearance and the same image size. We take the above considerations into account and suggest that the limitations of stimuli may matter for H1 but not H2-a.

4-5

Another reason why H2 but not H1 may be confirmed is that the Medusa effect has stronger consequences for implicit behavior, while explicit judgments are more variable; and viceversa, H1 but not H2 may be confirmed if the Medusa effect has stronger consequences for explicit (vs. implicit) behavior.

Reply:

Thank you so much for your thought-provoking suggestion. The explanation of this possible result was the part that we discussed significantly while writing the manuscript. We are excited to accept this suggestion and have added it in our manuscript based on your comment as follows:

[p.9]

Alternatively, the Medusa effect may have stronger consequences for implicit behavior.

4-6

While I empathize with the authors' preference for not conducting studies in the laboratory, I don't think this should be a determining factor in choosing which experiments to conduct (p. 7: "Since the COVID-19 pandemic is still in the process, we chose not to replicate them."), since of course this is an ongoing, constantly evolving situation. And beyond the online/inlab situation, I think the types of controls employed in Experiment 4 are important, and the authors should consider adopting those as well.

Reply:

Thank you so much for your understanding and suggestion. We apologize that we only mentioned the reason of COVID-19 but did not give further explanation as to why we chose not to conduct the original Experiment 4.

According to the original research, the purpose of Experiment 4 was to replicate the effect itself while equating facial appearance across levels of abstraction by employing new stimuli in which the same individuals appeared in L1 and L2, as well as matching image size across conditions. Furthermore, transitions between different levels of representation, which contained a new condition of L0 (a live face), were compared. In addition, stimuli of Experiment 5 were photographs used for Transition 2 in Experiment 4, in which facial appearance and image size were already equated.

However, the main point of the Medusa effect refers to the picture abstraction cost, which should be focused on the abstraction levels of pictures. Since the main purpose of our research is to examine the existence and generalization of the Medusa effect and its prosocial effects in Japan, we believe that the replication of Experiments 2 and 5 would be the best choice considering all the situations.

We have modified our manuscript to give a more detailed explanation as follows:

[p.7]

In reference to Experiment 4, Experiment 3 replicated the effect and further compared transitions by adding a new condition of a real person (L0). Since both Experiments 3 and 4 were conducted in the laboratory and the COVID-19 pandemic is still in process, it is hard for us to replicate them at this time. Furthermore, the Medusa effect itself refers to how the different abstraction levels of pictures affect mind perception. Consequently, we decided to not conduct Experiment 4 but only focus on the stimuli of pictures in our replication. Additionally, we will replicate Experiment 5 to explore whether different pictorial abstraction levels influence behavior in social interactions. Experiment 5 used the same pictures as Experiment 4 with better control of irrelevant variables (e.g., equated facial appearance across different abstraction levels and matched same image size) to conduct a mind perception task and the dictator game (Will et al., 2021).

4-7

For all studies, the authors need to specify a plan in case the main analyses aren't significant, to determine whether the results are null or inconclusive.

Reply:

Thank you for your comment. We are excited to accept your suggestion. We have added the equivalence test section in the data analysis, right after the main analysis part. If the results are not significant, we will then conduct the equivalence test to examine whether the non-significant results provide evidence for absence or a negligible effect size. We expect the equivalence test to further help us explain potential non-significant results. Furthermore, we will still conduct Study 3 (the conditional study) if the results of Studies 1 or 2 are non-significant, regardless of the result of the equivalence test. We have added this part to our revised manuscript as follows:

[p.14]

Equivalence test

If Study 1 does not replicate Will et al.'s (2021) Experiment 2 successfully, we will then conduct equivalence tests to examine whether the non-significant results provide evidence for the effect's absence or negligible size (Lakens et al., 2018).

The Smallest Effect Size Of Interest (SESOI) for our equivalence test was determined according to the small telescopes argument (Simonsohn, 2015) as the effect size the original design had 33% power to detect. For Realness, based on a power analysis and considering the 106 participants of the original experiment, a one-side paired t-test with an alpha of .05 would have had 33% power to detect an effect of $d = 0.1178$. This will be taken as our SESOI for Realness. Similarly, the SESOI will be $d = 0.1167$ for Agency and $d = 0.1195$ for Experience.

4-8

Minor points

Study 1:

The authors mention that "Similar to Will et al.'s (2021) study, pictorial abstraction is a between-subjects factor."; But abstraction is NOT a between-subjects factor in the original study ("Their task was to rate each of the two people shown in an image", p. 6). This should absolutely be corrected.

Reply:

Thank you so much for pointing this out. We apologize for such a careless mistake, and we have corrected it in our revised manuscript as follows:

[p.10]

Independent variable

Different levels of abstraction of people's photos. There are two abstraction levels in our study: L1 (picture of a person) and L2 (picture of a picture of a person). Similar to Will et al.'s (2021) study, pictorial abstraction is a within-subjects factor.

4-9

Instead, the original study varied the DV (Realness, Agency, and Experience) between subjects. This should absolutely be implemented.

Reply:

Thank you for pointing this out. We have added a detailed explanation of the conditions of the DV (Realness, Agency, and Experience) to the method part of Study 1 in our manuscript as follows:

[p.10]

Dependent variable

Mind perception. In Study 1, mind perception will be measured based on three dimensions, using the quantitative framework of mind perception that Will et al. (2021) used from 0 (lowest

level) to 10 (highest level). Participants will be randomly assigned into three groups, referring to the different tasks of assessing the Realness, Agency, and Experience of persons in L1 and L2.

4-10

Were the instructions directly translated from the original ones?

Reply:

Thank you for your comment. Yes, the instructions will be a Japanese-translated version of the original ones. We have added this detail to our revised manuscript as follows:

[p.12]

The participants will read the instruction (Japanese-translated version of the original one) and provide informed consent before participating in the study.

4-11

Will definitions also be provided as in the original study?

Reply:

Thank you for your comment. Yes, all the definitions will be provided as in the original study but in Japanese. We have added this detail into our revised manuscript as follows:

[p.13]

The question at the top will be about the person on the left side of the picture (e.g., Please rate the Experience (ability to feel) of the person on the left), and the one at the bottom will refer to the person on the right side of the screen (e.g., Please rate the Experience (ability to feel) of the person on the right).

4-12

This sentence confused me: "there will be no strict time limitation so that the participants can [...] take no longer than 5 minutes"

Reply:

Thank you for pointing this out. We are sorry for the confusing content. This sentence meant to express that there will be no strict time limitation for rating each picture, but the whole experiment will be completed in no longer than 5 minutes. Since the meaning of this sentence is quite confusing, we have removed the "there will be no strict time limitation so that" part from our manuscript.

4-13

The authors say they will recruit 564 participants, but the table then mentions "more than 564".

Reply:

Thank you for your comment. Based on the power analysis, the minimum sample size we need in Study 1 is 564 participants. We mentioned “more than 564” in the table, considering it was difficult for us to limit the number of participants to exactly 564 due to the characteristics of the participatory online recruitment system. We have added a detailed explanation of it in our revised manuscript as follows:

[p.12]

Considering it was difficult for us to limit the number of participants to exactly 564 due to the characteristics of the participatory online recruitment system, we will recruit at least 564 participants and use their data for the analysis based on the timestamp.

4-14

The table mentions "a paired t-test", but aren't the authors conducting three?

Reply:

Thank you for pointing this out, and we are sorry for the confusing content. The indefinite article “a” was used as we were referring to the analysis method itself, rather than the number of times we will conduct it. We have modified this in the revised table to avoid any misunderstanding as follows:

[p.22]

For H1, Study 1 uses three paired t-tests independently for each rating group to compare L1 and L2.

4-15

Again, I am not sure how the quality of the stimuli could explain failed replication of H1, if those are the same stimuli as in the original study.

Reply:

Thank you for your comment. Regarding the reasons, we have already provided an explanation above, please refer to 4-4.

4-16

Study 2:

What is the attention check?

Reply:

We apologize for not providing a clear explanation. The attention check is a procedure to ensure that participants completed the dictator game task in earnest. We have added the explanation in our revised manuscript as follows:

[p.17]

After the allocation is made, the participants will complete the final step as the attention check by selecting option four from a list of five options. Participants who failed the attention check will be excluded.

4-17

Why has the maximum donation amount (\$10) been lowered (to 1000 yen) when \$10 = ~1500 yen?

Reply:

Thank you for your comment. We did not set the maximum donation amount in real-time exchange rate but set it into a suitably round number according to daily-life experience.

4-18

Will et al. also had a rating phase in their Exp. 5; why is this omitted here?

Reply:

Thank you for pointing this out. It was omitted because we believed it was not necessary for us to perform it again at the beginning, since the differentiation between L1 and L2 of the stimuli has already been confirmed by the original Experiment 5. We have discussed it again and believed that it was our mistake for omitting it as prosocial behaviors could not be connected with mind perception without measuring it. We have added the two-alternative forced-choice (2AFC) task in our Study 2 as the original Experiment 5 in the first phase. We have modified it in our revised manuscript. See P.17 as follows:

[p.17]

Procedure

... ..

First, in the mind perception task, participants will be randomly shown one of four photographs, which depicts Person A and Person B. Participants will complete a two-alternative forced choice (2AFC) task by answering three questions: which person seems more real? Which person seems to have more Agency (ability to do)? Which person seems to have more Experience (ability to feel)?

4-19

p. 3: The authors define the Medusa effect as a tendency for people to 'evaluate a "picture of a person" as more mindful than a "picture of a picture of a person"', but people aren't rating the mindfulness of pictures (L1), they are rating the mindfulness of people (L0) in those pictures. This should of course be clarified.

Reply:

Thank you very much for pointing this out. We apologize for the incorrect definition, and we have rewritten it into the correct one in the manuscript as follows:

[p.3]

Will et al. (2021) found the tendency of people to evaluate a “person in picture” as more mindful than a “person in picture of a picture,” and named it “the Medusa effect.”

4-20

I found the writing often unclear, and I worry a naive reader unfamiliar with the Medusa effect might have a hard time following. I won't list all of them here, but I will just take the first few sentences from the abstract to exemplify:

The very first sentence "Pictures play an important role in containing and expressing information related to the human mind" is puzzling since many pictures are unrelated to the human mind (e.g. landscapes); do the authors mean pictures of people/faces?

4-21

The second sentence was also confusing "compositional differences [...] affect the way we perceive the vast amount of information" since it's unclear what the vast amount of information refers to; do the authors mean the way we perceive people?

4-22

The third sentence contains an incorrect definition of the Medusa effect, as per my point above.

4-23

The fourth sentence also confused me, since realness was never mentioned before (and is different from mindfulness; a rock can be real even though it doesn't have a mind); also, it's unclear what 'dimensions' refers to; do the authors mean 'abstraction' or 'compositionality' instead?

Reply:

Thank you for the comment. We have modified the abstract part based on your comments in our revised manuscript as follows. We hope the revised content will be easier for readers to understand.

[p.2]

With the extensive use of social media in the past decade, pictures of faces play an important role in containing and expressing information related to the human mind. Since pictures, as one of the features allowing them to carry a vast amount of information, can contain other pictures, compositional differences, such as picture abstraction levels (that is, a picture within a picture) also affect the way humans perceive the realness and mindfulness of the subjects depicted. However, this differential perception was often supposed to be on account of the impoverishment of a rich stimulus. Five past experiments found the tendency of people to evaluate a “person in picture” more mindful than a “person in picture of a picture” and named it “the Medusa effect.” This finding overturned the “impoverishment assumption” by suggesting that abstraction itself can reduce mind perception independent of stimulus richness, and it is critical for understanding the consequences of abstractions in perceiving and evaluating pictorial information in the world. Nevertheless, third parties are yet to replicate this study directly. Moreover, recent research has found a cross-culture effect on mind perception.

4-24

p. 5: "Following the aforementioned prior study, Will et al. (2021) used five experiments " I am not sure what the 'prior study' refers to.

Reply:

Thank you for the comment. The “prior study” refers to previous research on pictures and mind perception we have mentioned in the last paragraphs. To be specific, see [p.4/5]:

Besides this mythology story, researchers have found that pictures can convey more than the imagined. For example, pictures with eyes can attract our attention (Friesen & Kingstone, 1998). Moreover, an important feature of pictures caught the attention of the original researchers: they may not only contain partial information about reality but may also contain pictures that serve as different levels of abstraction (Will et al., 2021).

4-25

p. 5: I don't think eyetracking is considered physiological data?

Reply:

Thank you for your comment. We considered eye-tracking data as physiological data on account of regarding eye movement as a physiological process (which is a spontaneous process to some extent), although the result of the measurement could also be considered behavioral data related to the eyes. Since what we wanted to emphasize here is the finding of Experiment 3, which is that the attention bias between different levels occurs spontaneously during free viewing, we have modified the expression in our revised manuscript to avoid any misunderstanding as follows:

[p.7]

As for Experiment 3, eye-tracking data showed differentiation between L1 and L2 occurs spontaneously, and it explored the mechanism of the effect.