

**The Medusa effect: A registered replication report of
Will, Merritt, Jenkins, and Kingstone (2021)**

Jing Han^{1*}, Minjun Zhang¹, Jiaxin Liu¹, Yu Song¹, Yuki Yamada²

¹Graduate School of Human-Environment Studies, Kyushu University, Fukuoka,
Fukuoka, Japan

²Faculty of Arts and Science, Kyushu University, Fukuoka, Fukuoka, Japan

*Corresponding author:

E-mail: han.jing.888@s.kyushu-u.ac.jp (JH)

Abstract

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. Therefore, we plan to perform a replication of Will et al.’s (2021) Experiments 2 and 5 to examine the reliability, validity, and generalization of the Medusa effect in Japan, as well as its effect on prosocial behavior.

Keywords: Mind perception, The Medusa effect, Reality, Prosocial behavior, Dictator game

Introduction

The capacity to bridge the divide between one's own mind and others' minds is one of the most useful tools in social life. People use this impressive tool to understand, predict, and even control others' behavior and develop social connections with them (Waytz et al., 2010). During this process, most researchers have focused on how people perceive others' mental states, which is often referred to as "mind perception."

As pictures play an essential role in not only containing features but also expressing emotions, mind perception involving pictures is a theme worthy of research. Besides the apparent fact that the picture of a person would seem to have less realness than a person in reality for losing some features during pictorial abstraction. For example, a real person is more mindful than a portrait in a gallery. 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." This phenomenon is strikingly intriguing because it suggests that, when people evaluate the humanness of others, they may integrate information about the dimensions of the world, including their own. Specifically, the feature itself may considerably shape the perception, and the abstraction of information presentation is also an equally indispensable influencing factor which underlines the effect of dimensions the information presents.

The Medusa effect

Will et al. (2021) refer to the above psychological effect as the "Medusa effect," which is actually a reference to a Greek mythology story by the original authors. In mythology, Medusa was a snake-haired demoness in ancient Greece, who petrified anyone who saw her eyes. This demon was successfully decapitated by the hero Perseus with the assistance of Athena and Hermes, and it was through the shield-reflecting images that the power of her gaze was weakened, and Medusa was

successfully decapitated. While the fictional Medusa is L1 (“L” is short for level, referring to the levels of abstraction), Perseus sees Medusa on the reflective shield as L2. Moreover, the name "Medusa effect" better conveys that the higher the degree of abstraction carried by a picture, the weaker the power of mind perception conveyed. Besides this mythology story, researchers have found that pictures can convey more than [they](#) imagined. For example, pictures with eyes can attract our attention (Friesen & Kingstone, 1998). Furthermore, it has been reported (although reproducibility is debatable) that being gazed at [with](#) eyes in pictures increases prosocial behavior (Bateson et al., 2006). For instance, pictures of eyes can increase generosity in dictatorial games (Bardsley, 2008), charitable giving in the field (Ekström, 2012), as well as cooperative behavior in the field setting (Ernest-Jones et al., 2011). The mind perception thus conveyed, leads people to help or hurt others or to praise and punish others in their lives. Precisely because mind perception is the ability to reason about thoughts, numerous researchers have focused on how mind perception is defined, when it occurs, and its importance.

A considerable amount of research has proposed various frameworks for quantifying mind perception and its dimensions (Takahashi et al., 2016; Malle, 2019). However, the most influential one is the two-dimensional framework proposed by Gray et al. (2007), which refers to people intuitively thinking about other minds in terms of Experience (the capacity to sense and feel) and Agency (the capacity to plan and act). The original research on the Medusa effect adopted this framework and explored the differences in mind perception among different levels of pictorial abstraction.

Although mind perception occurs in the perceiver’s mind, the characteristics of the entity being perceived also influence mind perception (Waytz et al., 2010). Therefore, the original study focused on the effects of different abstractions on mind perception

of pictures. Pictures, as important bearers of expressions of people's ideas and perceptions of their authenticity, are uploaded billions of times daily on the Internet (Latha & Fathima, 2019). 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). In addition, the higher the abstraction level of a picture, the lower the potency of the subject. According to the above speculation, the original study primarily hypothesized that different levels of abstraction would bring about different levels of mind perception.

Following the aforementioned previous research, Will et al. (2021) used five experiments to verify the existence of the Medusa effect. First of all, it is necessary to clarify that the original study specified the real person as L0 (“L” is short for level, referring to the levels of abstraction), the picture of a person as L1, and the picture containing a picture of a person as L2, referring to different pictorial abstraction levels. Based on this setting, the original Experiments 1 and 2 tested the central hypothesis of the original study that L1 brings more mind perception than L2, by means of choice and evaluation. The original Experiment 3 conducted a free viewing task by using the same stimuli as Experiments 1 and 2 and monitored the looking behavior of participants via an eye tracker. The eye-tracking data showed that differentiation between L1 and L2 occurs spontaneously during passive viewing, even without an explicit mind perception task. Experiment 4 improved the stimuli to replicate the picture abstraction cost and further compared transitions between different levels of representation (L0 to L1, and L1 to L2). Finally, Experiment 5 used the dictator game to verify the possible effects of the Medusa effect on prosocial behavior by using the newly designed stimuli in Experiment 4. In previous research, dictator games were frequently used in various

studies investigating social norms, such as altruism and fairness (Guala & Mittone, 2010), or as an experimental method to measure pro-sociality. The main process of the game is that the participant controls a sum of money and decides how much money to give to the recipient; the remaining money goes to the participant. Furthermore, many studies have demonstrated that the amount of money distribution in the dictator game is significantly correlated with the social salience of the recipient (Guala & Mittone, 2010; Charness & Gneezy, 2008). [Experiment 5](#) linked behavior and cognition through the dictator game, verifying that picture abstraction would act on the game process through mind perception leading to [differences in allocation amounts](#).

[In conclusion, by conducting these five experiments, Will et al. \(2021\) first demonstrated the hypothesis that the abstraction level increases from L0 to L1 and from L1 to L2, thereby decreasing the level of mind perception associated with pictures. In addition, for the first time, the “Medusa effect” in mind perception was demonstrated, suggesting a new focus on psychological effects in a modern environment, where many emotions are communicated through the Internet and pictures. Moreover, mind perception plays a pivotal role in major areas of society such as education, law, medicine, and charity \(Gray, 2012\). Inevitably, it is necessary to consider the different behaviors resulting from varying degrees of mind perception through pictorial abstraction, which has attracted our attention.](#)

Aims of the present study

Will et al. (2021) conducted five experiments to determine the degree of mind perception between different levels of pictorial abstraction and its prosocial effects. For the first two experiments, we decided to replicate Experiment 2 rather than Experiment 1, to confirm the existence of the effect of using a rating task, which is more informative than the two-alternative forced choice (2AFC) task in Experiment 1 of the original

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

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.

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.

For the above reasons, in this research, we aim to replicate Experiments 2 and 5 of the original research, in Japan, to examine the existence and generalization of the Medusa effect and its prosocial effects. Moreover, in terms of sample size, 564 participants will be recruited for Study 1 and 660 will be recruited for Study 2, based on a prior power analysis.

Based on the above review, we plan to test the following hypotheses. For H1 (in Study 1), participants would rate L1 as having higher levels of both Realness, Agency, and Experience than L2. Study 2 will present three hypotheses, H2-a, H2-b and H2-c, which refer to the 2AFC task, the Dictator game, and the final individual difference analysis. Regarding H2-a, participants would choose L1 rather than L2 in all three Realness, Agency, and Experience dimensions. As for H2-b, in the dictator game, participants would allocate more money in condition L1 than in condition L2. With

respect to H2-c, the above analysis of individual differences in perception and behavior allows for the attribution of effects in the dictator game to the mind perception task. Accepting all three sub-hypotheses of Study 2 will be confirmed as support for H2 and the successful replication of Experiment 5.

If both H1 and H2 are supported, it confirms the original claim, with its reproducibility and generalizability extended. Picture abstraction cost and its behavior effect exist across different races of stimuli and participants. If H1 is supported, but H2 is not, the Medusa effect may exist but is not related to prosocial behaviors. We will explore the possible reason by conducting Study 3-b with the equated race between stimuli and participants. If H2 is supported, but H1 is not, the result may be caused by the limitations of the stimuli, since the stimuli of Study 2 are newly made presentations with better control of irrelevant variables, including equated facial appearance and the same image size. In contrast, the stimuli of Study 1 are pictures downloaded from the Internet that vary in terms of size, quality, gender, race, and emotional expression. Alternatively, the Medusa effect may have stronger consequences for implicit behavior. If neither H1 nor H2 are supported, the reproducibility of the Medusa effect fails. We will redesign the stimuli with the same ethnicity as the participants and further attempt additional replications to explore possible reasons for the failure.

We plan to conduct the conditional study (Study 3) only if we have not replicated Studies 1 or 2 successfully, by using newly designed stimuli with higher quality and the same ethnicity of participants to further replicate the original Experiments 2 or 5. In case H1 is not supported, we will conduct Study 3-a to test H1 again. If Study 2 fails to replicate the original Experiment 5, we will conduct Study 3-b, hoping to conduct a further replication with newly constructed stimuli. If both Studies 1 and 2 fail to

replicate the original experiments, we will use improved stimuli to conduct additional replication, referring to Study 3 (a and b), further examining the Medusa effect.

Study 1

Study 1 employs the mind perception assessment task to examine whether pictorial abstraction levels differ in the degree of mind perception, which is a direct replication of Will et al.'s (2021) Experiment 2, to examine the existence of the Medusa effect in Japan.

Method

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.

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.

Participants

Sample size and power analysis

[At least](#) 564 participants will be recruited for Study 1 based on a prior power analysis using G*Power 3.1.9.7 for Windows (Faul et al., 2007, 2009) to replicate the Medusa effect of the study by Will et al. (2021). In their study, 320 participants were randomly assigned into three groups to rate Realness (n = 107), Experience (n = 109),

and Agency ($n = 104$) of the pictures. Three paired t -tests were conducted independently for each rating group to compare L1 and L2. The results of the analysis revealed that all the groups, Realness, Experience, and Agency, of L1 compared with L2 reached significant differences (all $ps < .05$), and for each of the effect size, Cohen's d was given ($d_R = 0.83$, $d_E = 0.34$, $d_A = 0.39$). However, the findings and statistical results in their study are the latest research developments, and other reference studies are rarely available. In addition, the effect size can sometimes be overestimated owing to the small sample size and tends to decrease in subsequent replication studies, which is a statistical bias named "Winner's Curse" (Button et al., 2013; Nitta et al., 2018; Yonemitsu et al., 2020). Based on these two points, we plan to use a small effect size ($d = 0.2$), which was defined by Cohen (1969). We conducted a one-tailed, paired t -test power analysis by assuming Cohen's $d = 0.2$ as small effect size, significance level $\alpha = .05$, and power level $1 - \beta = .80$ (Cohen, 1988) to calculate our sample size. The results indicated that 156 participants will be required per group (i.e., 468 in total). In addition, it is conceivable that a certain number of participants may withdraw from the experiment midway due to dissatisfaction; hence, we added approximately 20% to this number in case of power loss due to data exclusions (i.e., 562 in total). To equalize the number of recruitments for the three groups, we added two to this number (i.e., 564 in total).

Recruitment and screening

We will recruit participants in Japan via the Yahoo! Crowdsourcing Service (<https://crowdsourcing.yahoo.co.jp/>). All participants will complete the study online in exchange for monetary compensation. All questions need to be filled out for the participants to submit the answer, and we will screen the participants' IP addresses to preclude repeat submissions. Data will continue to be collected until a minimum

sample size of 564 is reached, as indicated by the sample size analysis. Considering that 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.

Stimuli and design

The original stimuli (pictures) were provided to us via email by the original authors on request. We will use the same stimuli as in the original research. The stimuli contain a total of 29 pictures, and each picture depicts an L1 person and an L2 person, which present different degrees of abstraction in a single scene. For example, a person (L1) holding a portrait photograph (L2), or a computer user (L1) with an onscreen interlocutor (L2). The depicted L1 and L2 of each scene will vary in terms of size, quality, and on-screen location (left or right), and the depicted L1 and L2 persons will also vary in age, gender, race, and emotional expression. All the scenes are cropped to a standard size of 400 pixels high \times 600 pixels wide.

Pictorial abstraction (L1 and L2) is a *within*-subjects factor. Participants will be randomly assigned to *one* of the Experience, Agency, or Realness conditions, which refer to their evaluation task of both L1 and L2, to examine whether there is a difference in mind perception between the different pictorial abstractions.

Procedure

The participants will read *the instruction (Japanese-translated version of the original one)* and provide informed consent before participating in the study. They will also be informed that they can withdraw participation at any time. We will not collect any personal information except for gender and age. The collected data will be strictly protected.

Our study will strictly follow the same procedures used in the study by Will et al. (2021), except for using the Yahoo! Crowdsourcing Service to recruit participants in Japan and presenting the instructions and questions in Japanese.

At the beginning of the experiment, demographic information on the participants' age and gender will be collected. Thereafter, the participants will be shown pictures consisting of two people with different abstraction levels (L1 and L2). Their task will be to rate each of the two people based on Experience, Agency, or Realness. For each trial, a single picture will be shown to the participants on the screen with attribute questions (e.g., Experience) above and below it. 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). Participants will move a slider to a whole number on a scale ranging from 0 (the lowest level) to 10 (the highest level) to answer each question. The trial order is randomized, and the participants can complete the experiment at their own pace, but take no longer than 5 minutes.

Data analysis

Main analysis

Since this study is a replication of the Medusa effect, we will analyze the data in the same way as Will et al. (2021) did. **We** will compare whether there are significant differences between participants' perceptions of L1 and L2 on the three dimensions of Realness, Agency, and Experience based on paired *t*-tests. Confirmation of our hypothesis is based on the following criteria.

For H1, we predict that participants will rate L1 (picture of a person) as having higher levels of Realness, Agency, and Experience than L2 (picture of a picture of a

person). Significantly higher scores ($\alpha = .05$) for L1 than L2 will indicate acceptance of H1 as well as successful replication of Will et al.'s (2021) Experiment 2.

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.

Study 2

Study 2 employs a **mind perception task** and a dictator game task to examine whether pictorial abstraction levels affect conduct in social interactions, which is a replication of Will et al.'s (2021) Experiment 5, in Japan.

Method

Independent variable

Different levels of abstraction of people's photos. Our study has two abstraction levels: L1 and L2. Similar to Will et al.'s (2021) study, pictorial abstraction is a within-subjects factor.

Levels of abstraction of recipients in the dictator game. There are two levels of abstraction for the receiver *in the dictator game*, classified as L1 or L2. This is a *between-subjects factor*.

Dependent variable

Mind perception. In Study 2, mind perception will be measured through a two-alternative forced choice (2AFC) task that indicates which of the two individuals (L1 or L2) has higher attributes of Realness, Agency, and Experience.

Amount of allocated money in the dictator game. In the dictator game, participants will decide the amount of money they would like to allocate from 0 to 1,000 Japanese yen to L1 person or L2 person.

Participants

Sample size and power analysis

In Study 2, we obtained our sample size based on the given effect size Cohen's d ($d = 0.36$) of the Wilcoxon-Mann-Whitney test in the original study, similar to Study 1. In line with the same principle to avoid possible existence bias in the replication study, we also used a small effect size d value and conducted a non-parametric independent sample power analysis of one-tailed, normal parent distribution, assuming an effect size Cohen's $d = 0.2$, significance level $\alpha = .05$, power level $1 - \beta = .80$, to compute the sample size. In addition, considering the necessity of counterbalance, we decided to equalize the participants of the two groups by setting the allocation ratio $N2/N1$ to 1. The results revealed that 325 participants per group were required to obtain a power of .80. Essentially, there is no case of data not being collected by participant withdrawal since the agreement of participants granted for only one assessment-allocation task in Study 2. Nevertheless, 0.5% of the data loss occurred due to unknown failure in the original study, and approximately 1% of the data were excluded because

of failed attention checks. Thus, we added an additional 1.5%, which is five people to each group in Study 2, and **at least** 660 participants (330 per group) will be recruited in total.

Recruitment and screening

Study 2 will use the same recruitment and screening methods as Study 1, except that participants who had completed Experiment 1 will be excluded. The minimum sample size of Study 2 is 660 participants. **Considering that it was difficult for us to limit the number of participants to exactly 660 due to the characteristics of the participatory online recruitment system, we will recruit at least 660 participants and use their data for the analysis based on the timestamp.**

Stimuli and design

The original stimuli (pictures) have already been provided to us via email by the original authors. We will employ the same stimuli as in the initial study. Four photographs of two experimental model volunteers (Person A and Person B) will be presented as stimuli. These photographs consist of two versions (original displays and horizontally inverted mirror displays) of two original photographs, balancing the levels of pictorial abstraction (L1, L2) and spatial location (left, right). For the original photographs, each captures the entire face of one model together with the life-sized photo of another model that she is holding. For example, the photograph depicts Person A (L1) holding a life-sized photo of Person B (L2), or Person B (L1) holding a life-sized photo of Person A (L2). Each of the four pictures has been cropped to 1800 pixels high by 2400 pixels wide, to be displayed on the screen during the experiment.

Following the original design of Will et al. (2021), **participants** will be randomly assigned to **four versions** of **display (photograph)** and **complete the mind perception task**, which is to **decide which of two people (person A or person B) seems to have**

higher Realness, Agency, and Experience. Afterward, the dictator game will be conducted based on the same display, in which participants will be randomly assigned the L1 person or the L2 person as the recipient, to examine the connection between pictorial abstraction and prosocial behavior. Finally, we will conduct an attention check to ensure that the data are valid.

Procedure

Study 2 will strictly follow the same procedures used in Experiment 5 by Will et al. (2021). Participants will be recruited in the same criteria, and the experimental statement administered to participants will be the same as in Study 1.

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)?

Afterward, participants will proceed to a one-shot dictator game using the same display. They will be randomly assigned to groups where L1 or L2 as the designated recipient, indicated by an onscreen arrow and text instructions. The task is to share 1,000 Japanese yen endowment with a specified onscreen recipient. A slider in yen (0–1,000) will be displayed at the bottom of the screen, and participants will manipulate the slider to decide the amount of allocation to the recipient. 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.

Data analysis

Since Study 2 is a replication of Will et al. 's (2021) Experiment 5, we will analyze the data of the [mind perception task and the dictator game](#) in the same manner as Will et al. (2021) did in their Experiment 5. [For the mind perception task, we will conduct a Binomial test to compare the proportion of participants choosing L1 and L2.](#) [Subsequently, a Mann-Whitney test will be conducted to compare money allocations between L1 recipients and L2 recipients in all four counterbalanced versions of the stimuli.](#) [Finally, we will conduct a Mann-Whitney test to compare whether participants who perceived L1 as higher than L2 on all three dimensions in the mind perception task made a strong distinction between L1 and L2 in the money allocation of the dictator game.](#)

For H2-a, we predict that [the proportion of participants choosing L1 will be higher than that of participants choosing L2 for all three dimensions \(Realness, Agency, and Experience\) in the mind perception task.](#) [A proportion that is significantly above \(\$\alpha = .05\$ \) the chance level of 50% will indicate acceptance of H2-a.](#) Moreover, for H2-b, [significantly more money \(\$\alpha = .05\$ \) allocated to L1 recipients than L2 recipients will indicate acceptance of H2-b.](#) [As for H2-c, participants who perceived L1 as higher than L2 on all three dimensions and differentiated significantly between L1 and L2 \(\$\alpha = .05\$ \) in their dictator game allocations will indicate acceptance of it.](#) [Acceptance for all three hypotheses of Study 2 will indicate the success of the replication of Will et al.'s \(2021\) dictator game task in Experiment 5.](#)

Study 3 (Conditional Study)

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.

As for redesigning the stimuli, we would consider the factor of cross-cultural differences and further enhance the control of irrelevant variables. The original stimuli were pictures of Western people, which were different from our participants in Japan. Certain factors in the visual cognition of faces of different ethnicities (e.g., overall deterioration of identification/discrimination) could be the reason that the replication fails. Therefore, we would redesign the stimuli using Japanese faces. Moreover, instead of the pictures downloaded from the Internet in the original research, we would take pictures of volunteers in controlled conditions ourselves to exclude irrelevant variables, including, but not limited to, gender, expression, size, and angle of the portraits.

Study 3-a

Study 3-a is a replication of Will et al.'s (2021) Experiment 2 with newly created stimuli. If H1 is not supported in Study 1, we will conduct Study 3-a to further examine the existence of the Medusa effect. We hypothesize that participants would rate L1 as having higher levels of Realness, Agency, and Experience than L2.

Method

Stimuli, design, participants, and procedure

Similar to Study 1, the stimuli of Study 3-a will be 30 pictures; each picture depicts an L1 person and an L2 person, which present different degrees of abstraction in a

single scene. For example, one person (L1) holds a portrait photograph (L2). We will design new stimuli by recruiting Japanese volunteers to capture their portraits of the entire face. Irrelevant variables, including but not limited to expression, size, and angle, will be controlled while taking the pictures. Gender will be counterbalanced in redesigned stimuli.

The design, sample size, recruitment and screening, and procedure of Study 3-a will be the same as those of Study 1.

Study 3-b

Study 3-b is a replication of Will et al.'s (2021) Experiment 5, except for the use of newly designed stimuli. If our Study 2 fails to replicate the original Experiment 5, Study 3-b will be conducted. The hypotheses of Study 3-b are the same as those of our Study 2.

Method

Stimuli, design, participants, and procedure

Similar to Study 2, the stimuli in Study 3-b will be four photographs of two Japanese model volunteers (Person A and Person B) recruited in Japan. Each of these photographs captures the entire face of one model together with the life-sized photo of another model that she is holding. The photographs consist of two versions (original displays and horizontally inverted mirror displays) of the two original photographs, balancing the pictorial abstraction levels (L1, L2) and spatial location (left, right). For example, the photograph depicts Person A (L1) holding a life-sized photo of Person B (L2) or Person B (L1) holding a life-sized photo of Person A (L2). Irrelevant variables, including but not limited to expression, size, and angle, will be controlled while taking the new pictures.

The design, sample size, recruitment and screening, and procedure of Study 3-b will be the same as that of Study 2.

Ethics

The current study has been committed by the ethics committee of Kyushu University (Protocol Number: 2022-015), and will be conducted in accordance with the principles of the Declaration of Helsinki. The participants will have the right to self-determine to cease the study at any time without any disadvantages. All participants will be provided informed consent, and the study will be started only with the granted approval. The personal information of all participants will be strictly protected and will not be disclosed to third parties.

Question	Hypothesis	Sampling plan	Analysis Plan	Rationale for deciding the sensitivity of the test for confirming or disconfirming the hypothesis	Interpretation, given the different outcomes
Q1: Does the level of pictorial abstraction affect mind perception?	H1: Participants would rate L1 (person in picture) as having higher levels of Realness, Agency, and Experience than L2 (person in picture of a picture).	564 participants will be recruited in Study 1. The number of participants is based on a power analysis.	Same as the original study; for H1, Study 1 uses three paired <i>t</i> -tests independently for each rating group to compare L1 and L2.	Significantly higher scores in Realness, Agency, and Experience of L1 as compared to those of L2 ($\alpha = .05$) would indicate the acceptance of H1 and the successful replication of Will et al.'s (2021) Experiment 2.	If H1 is not supported, the replication of the Medusa effect fails. The results may be caused by the limitations of the stimuli (e.g., quality and race). We will redesign the stimuli with higher quality and the same ethnicity as the participants, and further try additional replications (Study 3-a).
Q2-1: Does the level of pictorial abstraction affect mind perception (with stimuli	H2-a: Participants would perceive L1 to be higher than L2 in all three dimensions (Realness,	660 participants will be recruited in Study 2. The number of participants is	Same as the original study; for H2-a, we will use a Binomial test to compare the proportion of participants	Significantly higher ($\alpha = .05$) proportion of participants choosing L1 over L2 than the chance level of 50% will indicate acceptance of H2-a.	If H2 is not supported, it would suggest that the Medusa effect is not related to prosocial behaviors. We will redesign the stimuli with the same ethnicity as the participants, and further try

<p>that have better control of irrelevant variables)?</p>	<p>Agency, and Experience).</p>	<p>based on a power analysis.</p>	<p>choosing L1 and L2.</p>		<p>additional replications (Study 3-b).</p>
<p>Q2-2: Does the level of pictorial abstraction affect conduct in social interaction?</p>	<p>H2-b: In the dictator game task, participants would allocate more money to L1 recipients than L2 recipients.</p>		<p>Same as the original study; for H2-b, we will conduct a Mann-Whitney test to analyze the data of the dictator game task.</p>	<p>Significantly more money ($\alpha = .05$) allocated to L1 recipients than L2 recipients would indicate the acceptance of H2-b.</p>	
<p>Q2-3: Can the effects in the dictator game be attributed to effects in the mind perception task?</p>	<p>H2-c: Mind perception distinction would affect conduct in the dictator game.</p>		<p>Same as the original study; for H2-c, we will conduct a Mann-Whitney test to compare participants' individual differences in perception and behavior.</p>	<p>Participants who perceived L1 as higher than L2 on all three dimensions and differentiated significantly between L1 and L2 ($\alpha = .05$) in their dictator game allocations will indicate acceptance of H2-c.</p>	

<p>Q3-1: Does the level of pictorial abstraction affect mind perception (with newly made stimuli)?</p>	<p>H3-1: Participants would rate L1 as having higher levels of Realness, Agency, and Experience than L2.</p>	<p>Same as our Study 1.</p>	<p>Same as our Study 1; for H3-1, Study 3-a uses tree paired t-tests independently for each rating group to compare L1 and L2.</p>	<p>Significantly higher scores in both Realness, Agency, and Experience of L1 than scores of L2 ($\alpha = .05$) would indicate the acceptance of H3-1 and the successful replication of Will et al.'s (2021) Experiment 2.</p>	<p>If H3-1 is supported, it would suggest that the Medusa effect exists, and the race of stimuli may be the reason for the unsupported H1. If H3-1 is not supported, it would suggest that there is a possibility that the Medusa effect does not exist, or only exists under very limited conditions.</p>
<p>Q3-2a: Does the level of pictorial abstraction affect mind perception (with newly made stimuli)?</p>	<p>H3-2a: Participants would perceive L1 to be higher than L2 in all three dimensions (Realness, Agency, and Experience).</p>	<p>Same as our Study 2.</p>	<p>Same as our Study 2; for H3-2a, we will use a Binomial test.</p>	<p>Significantly higher ($\alpha = .05$) proportion of participants choosing L1 over L2 than the chance level of 50% will indicate acceptance of H3-2a.</p>	<p>If H3-2 is supported, it would suggest that the Medusa effect affects prosocial behaviors, the race of stimuli may be the reason for the unsupported H2. If H3-2 is not supported, it would suggest that the Medusa effect is not related to prosocial behaviors.</p>
<p>Q3-2b: Does the level of pictorial abstraction affect conduct in social interaction</p>	<p>H3-2b: In the dictator game task, participants would allocate more money to L1 recipients than L2 recipients.</p>		<p>Same as our Study 2; for H3-2b, we will use a Mann-Whitney test.</p>	<p>Significantly more money allocated to L1 recipients than L2 recipients ($\alpha = .05$) would indicate the acceptance of H3-2b.</p>	

(with newly made stimuli)?			
Q3-2c: Can the effects in the dictator game be attributed to effects in the mind perception task? (with newly made stimuli)	H3-2c: Mind perception distinction would affect conduct in the dictator game.	Same as our Study 2; for H3-2c, we will use a Mann-Whitney test.	Participants who perceived L1 as higher than L2 on all three dimensions and differentiated significantly between L1 and L2 ($\alpha = .05$) in their dictator game allocations will indicate acceptance of H3-2c.

References

- Bardsley, N. (2008). Dictator game giving: Altruism or artifact? *Experimental Economics*, *11*(2), 122-133. <https://doi.org/10.1007/s10683-007-9172-2>
- Bateson, M., Nettle, D., & Roberts, G. (2006). Cues of being watched enhance cooperation in a real-world setting. *Biology Letters*, *2*(3), 412-414. <https://doi.org/10.1098/rsbl.2006.0509>
- Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, *14*(5), 365–376. <https://doi.org/10.1038/nrn3475>
- Charness, G., & Gneezy, U. (2008). What's in a name? Anonymity and social distance in dictator and ultimatum games. *Journal of Economic Behavior & Organization*, *68*(1), 29-35. <https://doi.org/10.1016/j.jebo.2008.03.001>
- Cohen, J. (1969). *Statistical Power Analysis for the Behavioral Sciences*. New York: Academic Press.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences (2nd ed.)*. New York: Academic Press. <https://doi.org/10.4324/9780203771587>
- Ekström, M. (2012). Do watching eyes affect charitable giving? Evidence from a field experiment. *Experimental Economics*, *15*(3), 530-546. <https://doi.org/10.1007/s10683-011-9312-6>
- Ernest-Jones, M., Nettle, D., & Bateson, M. (2011). Effects of eye images on everyday cooperative behavior: A field experiment. *Evolution and Human Behavior*, *32*(3), 172-178. <https://doi.org/10.1016/j.evolhumbehav.2010.10.006>

- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. <https://doi.org/10.3758/BRM.41.4.1149>
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. <https://doi.org/10.3758/BF03193146>
- Friesen, C. K., & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin & Review*, *5*(3), 490-495. <https://doi.org/10.3758/BF03208827>
- Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychological bulletin*, *133*(4), 694. <https://doi.org/10.1037/0033-2909.133.4.694>
- Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science*, *315*(5812), 619-619. <https://doi.org/10.1126/science.1134475>
- Gray, K., Young, L., & Waytz, A. (2012). Mind perception is the essence of morality. *Psychological Inquiry*, *23*(2), 101-124. <https://doi.org/10.1080/1047840X.2012.651387>
- Guala, F., & Mittone, L. (2010). Paradigmatic experiments: The dictator game. *The Journal of Socio-Economics*, *39*(5), 578-584. <https://doi.org/10.1016/j.socec.2009.05.007>
- Krumhuber, E. G., Swiderska, A., Tsankova, E., Kamble, S. V., & Kappas, A. (2015). Real or artificial? Intergroup biases in mind perception in a cross-cultural perspective. *PLoS One*, *10*(9), e0137840. <https://doi.org/10.1371/journal.pone.0137840>

- Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science*, 1(2), 259-269. <https://doi.org/10.1177/2515245918770963>
- Latha, P. M., & Fathima, A. A. (2019). Collective compression of images using averaging and transform coding. *Measurement*, 135, 795-805. <https://doi.org/10.1016/j.measurement.2018.12.035>
- Malle, B. (2019). How many dimensions of mind perception really are there? *Proceedings of the 41st Annual Meeting of the Cognitive Science Society* (pp. 2268-2274).
- Nitta, H., Tomita, H., Zhang, Y., Zhou, X., & Yamada, Y. (2018). Disgust and the rubber hand illusion: A registered replication report of Jalal, Krishnakumar, and Ramachandran (2015). *Cognitive Research: Principles and Implications*, 3(1), 15. <https://doi.org/10.1186/s41235-018-0101-z>
- Simonsohn, U. (2015). Small telescopes: Detectability and the evaluation of replication results. *Psychological science*, 26(5), 559-569. <https://doi.org/10.1177/0956797614567341>
- Takahashi, H., Ban, M., & Asada, M. (2016). Semantic differential scale method can reveal multi-dimensional aspects of mind perception. *Frontiers in Psychology*, 7, 1717. <https://doi.org/10.3389/fpsyg.2016.01717>
- Waytz, A., Gray, K., Epley, N., & Wegner, D. M. (2010). Causes and consequences of mind perception. *Trends in Cognitive Sciences*, 14(8), 383-388. <https://doi.org/10.1016/j.tics.2010.05.006>

Will, P., Merritt, E., Jenkins, R., & Kingstone, A. (2021). The Medusa effect reveals levels of mind perception in pictures. *Proceedings of the National Academy of Sciences*, *118*(32), e2106640118. <https://doi.org/10.1073/pnas.2106640118>

Yonemitsu, F., Ikeda, A., Yoshimura, N., Takashima, K., Mori, Y., Sasaki, K., Qian, K., & Yamada, Y. (2020). Warning 'Don't spread' versus 'Don't be a spreader' to prevent the COVID-19 pandemic: Preventing messages in COVID-19 pandemic. *Royal Society Open Science*, *7*(9), [200793]. <https://doi.org/10.1098/rsos.200793>

Acknowledgements

We would like to thank Editage (www.editage.com) for the English language editing.

Authors' contributions

CRedit Statement for Stage 1 manuscript

Conceptualization: JH, MZ, JL, YS, and YY; Funding acquisition: YY; Methodology: JH, MZ, JL, YS, and YY; Project administration: YY; Supervision: YY; Writing - original draft: JH, MZ, JL, YS, and YY

Funding

This research is supported by JSPS KAKENHI: 20H04581, 21H03784, and 22K18263.