Two replications of "Hierarchical encoding makes individuals in a group seem more attractive (2014; Experiment 4)"

Yuko Ojiro ✉, a, Akihiko Gobara ✉,a, Giyeon Nam a, Kyoshiro Sasaki a,b, Reiki Kishimoto a,c, Yuki Yamada d, Kayo Miura e

a Graduate School of Human-Environment Studies, Kyushu University

b Japan Society for the Promotion of Science

c Graduate School of Letters, Kyoto University

d Faculty of Arts and Science, Kyushu University

e Faculty of Human-Environment Studies, Kyushu University

Abstract ▪ The cheerleader effect implies that a person in a group look like more attractive than in isolation. Walker and Vul (2014) reported results supporting the existence of the cheerleader effect. We replicated Walker and Vul's Experiment 4, which manipulated group size. Their participants were asked to rate attractiveness of each female face image in a group (one of 4, 9, or 16 members) and in isolation and revealed that attractiveness ratings significantly increased in all the group conditions. We performed two direct replications of this experiment using Japanese participants. As a result, at least one experiment yielded the pattern of the results similar to those of the previous study, although the effect was not significant and the effect size was small.

Keywords ▪ Cheerleader effect; replication study

✉ ojryuko@gmail.com ackey0299@gmail.com

Introduction

A girl looks more attractive when she is in a group than when she is alone (*the cheerleader effect*). The expression "cheerleader effect" was coined in the TV sitcom "*How I Met Your Mother*" (Rashid & Fryman, 2008). Walker and Vul (2014) tested whether the cheerleader effect really exists by presenting an individual face image in a group or in isolation and showed that the face was evaluated as more attractive in a group than in isolation. Their findings suggested that the cheerleader effect could occur in laboratory.

Straightforward experimental evidence was found in their Experiment 4. Thirty-seven undergraduate students (13 men, 24 women) from the University of California, San Diego participated in the experiment. The researchers varied the number of faces (from 0 to 16) surrounding a target face to be rated. As a result, attractiveness ratings with the surrounding faces were significantly higher than those with no surrounding face, although there was no effect of the number of surrounding faces. They suggested the ensemble coding mechanism in the visual system, especially for processing facial parts, involves the cheerleader effect (see also van Osch, Blanken, Meijs, & van Wolferen, 2015). Therefore, researchers in this field should test

this ensemble-coding hypothesis, and that is why we first tested the robustness of this phenomenon. Thus, we directly replicated their Experiment 4 as a first step.

Experiment 1

Method

Participants

Ninety-five Japanese people (64 men, 31 women) participated in this experiment (mean age = 39.8, SD = 8.98) for 28 JPY cash payment. All participants were naive as to the aim of this experiment.

Stimuli and Procedure

Stimuli were identical to those used in the previous study (Walker & Vul, 2014), which were provided by Walker and Vul in a personal communication. The size of each image was 118 × 129 pixels. As in the previous study, we randomly chose 77 faces as target faces. The remaining 223 face images were used as flanker faces in the group conditions. Images used as the target faces were randomly chosen across participants. There were four group size conditions in the Experiment 1: the 1
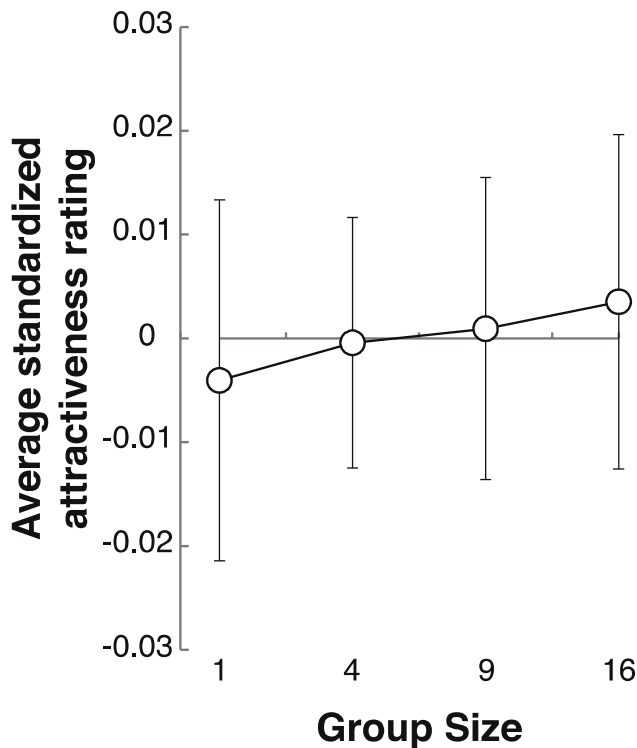
Figure 1 ▪ Mean standardized attractiveness scores in each group size in Experiment 1. Error bars represent 95% confidence intervals.

Figure 2 ▪ Mean standardized attractiveness scores in each group size in Experiment 2. Error bars represent 95% confidence intervals.

face (i.e., isolated), 4 faces, 9 faces, and 16 faces conditions. Each target was presented once in each of the four conditions. Faces were presented within each cell on imaginary square grids (1 × 1 for the 1 face condition, 2 × 2 for the 4 faces condition, 3 × 3 for the 9 faces condition, and 4 × 4 for the 16 faces condition; Figure 1). In the 4, 9, and 16 faces conditions, the stimuli consisted of one target face and flankers. The procedure was the same as that of the previous study. The present experiments were conducted online. The participants were asked to rate attractiveness of the target face. Each participant performed 308 trials: four group size conditions × 77 target faces. The trial order was randomized across the participants.

Results and discussion

We conducted a one-way within-group analysis of variance (ANOVA) with group size as a factor. Using the method exactly same as Experiment 4 of Walker and Vul (2014), we z-transformed rating scores to factor out between-participants variation of scale usage and variability in the actual attractiveness of a given face. Figure 1 shows the mean z-scores in each condition in
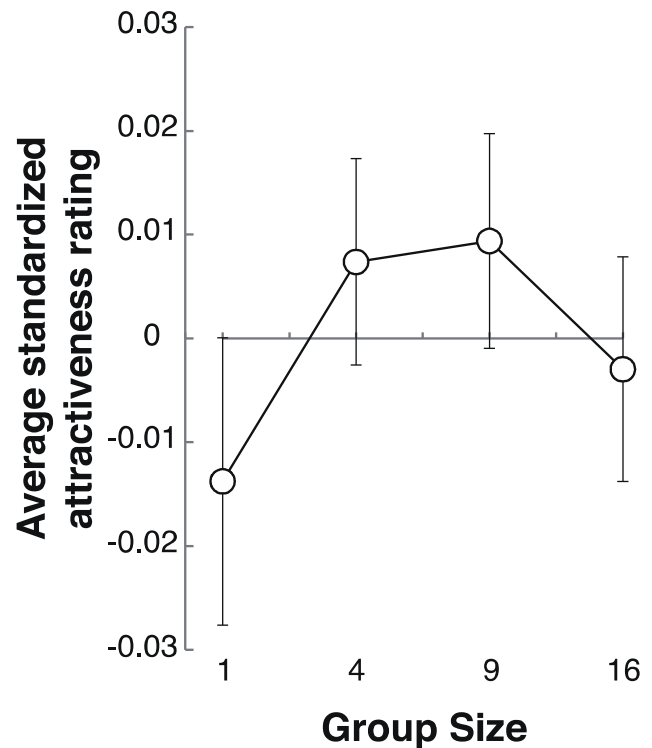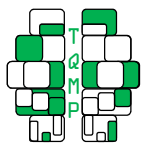
this experiment. Assumption of sphericity was violated, $\chi^2(5) = 38.17$, $p < .0001$, and hence, we adjusted degrees of freedom with Greenhouse-Geisser's Epsilon ($\varepsilon = .773$). The ANOVA showed no significant main effect of group size, $F(2.32, 218.01) = 0.128$, $p = .91$, $\eta_p^2 = .001$. Considering the non-significant main effect and the very small effect size, the results suggest that we could not replicate the cheerleader effect.

Experiment 2

The present Experiment 1 failed to replicate the results of Walker and Vul's Experiment 4 using the same stimuli as the previous study. It is highly likely that American people participated in the experiments of the previous study. This means that in the previous study American people evaluated faces of people in their own country. It is well known that own-race and other-race faces are processed differently (Golby, Gabrieli, Chiao, & Eberhardt, 2001; Young, Hugenberg, Bernstein, & Sacco, 2012). Thus, we believed that a truly direct replication is the situation in which Japanese people evaluated Japanese faces. Therefore, we re-ran the present first experiment by employing 300 newly

introduced Japanese faces.

Method

Experiment 2 was identical to Experiment 1 except that the stimuli were 300 Japanese facial images collected from the Web. The pictures used in this experiment can be found on the journal's web site. One hundred and fifty-three Japanese people (98 men, 55 women) were newly recruited for this experiment (mean age = 39.99, SD = 9.92).

Results and discussion

Figure 2 shows the results of Experiment 2. Assumption of sphericity was violated, $\chi^2(5) = 26.70$, $p < .0002$, and hence, we adjusted degrees of freedom with Greenhouse-Geisser's Epsilon ($\varepsilon = .887$). A one-way ANOVA showed no significant main effect of group size on attractive ratings, $F(2.66, 404.55) = 2.587$, $p = .060$, $\eta_p^2 = .017$. Although the main effect nearly reached significance, the effect size was quite small (the η_p^2 value in Experiment 4 of the original study was .197). Thus, in terms of p values, we again failed to replicate the results of Walker and Vul's Experiment 4, even when we used Japanese facial stimuli. However, the pattern of the results was quite similar to the one of the previous study.

General discussion

The present study replicated Walker and Vul's Experiment 4 twice, and the results showed that the main effects did not reach significance. Here we discuss factors that might have produced the discrepancy between the present study and previous study. First, it is possible that the sample size affected the significance of frequency-based statistics like ANOVA used in both the studies. While the sample sizes in the present experiments were 2.5 and 4.1 times as large as that of previous study, we did not obtain significant main effects. This thing suggests that the sample size in the present experiments did not prevent significant main effects. The very small effect sizes also suggest no cheerleader effect in the present study. Second, we used only Japanese people as participants. Some sort of cultural differences in face perception or cognition might have affected the results, although there is no study showing cultural differences in the visual ensemble coding of facial parts to the best of our knowledge. Like this, the cheerleader effect appears difficult to replicate in a wide range of settings at least

for Japanese observers. On the other hand, we found that in Experiment 2 the pattern of the results was strikingly similar to that in the previous study, suggesting that consistency in race between stimuli and observers may be important to producing the cheerleader effect. To find out the range of settings that can produce the cheerleader effect, further replication studies will be required before starting to investigate the underlying mechanism.

Authors' notes and acknowledgments

References

Golby, A. J., Gabrieli, J., Chiao, J. Y., & Eberhardt, J. L. (2001). Differential responses in the fusiform region to same-race and other-race faces. *Nature Neuroscience*, *4*, 845–850.

Rashid, R., (Writer), & Fryman. P. (Director). (2008). Not a father's day [Television series episode]. In C. Bays & C. Thomas (Creators), *How I Met Your Mother*, New York, NY: CBS.

van Osch, Y., Blanken, I., Meijs, M. H. j., & van Wolferen, J. (2015). A group's physical attractiveness is greater than the average attractiveness of its members: The group attractiveness effect. *Personality and Social Psychology Bulletin*, *41*, 559-74.

Walker, D., & Vul, E. (2014). Hierarchical encoding makes individuals in a group seem more attractive. *Psychological Science*, *25*, 230-235.

Young, S. G., Hugenberg, K., Bernstein, M. J., & Sacco, D. F. (2012). Perception and motivation in face recognition: a critical review of theories of the Cross-Race Effect. *Personality and Social Psychology Review*, *16*, 116–142.
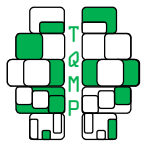
Citation

Ojiro, Y., Gobara, A., Nam, G., Sasaki, K., Kishimoto, R., Yamada, Y., & Miura, K. (2015). Two replications of "Hierarchical encoding makes individuals in a group seem more attractive (2014; Experiment 4)". *The Quantitative Methods for Psychology*, *11* (2), r8-r11.