



Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

journal homepage: www.elsevier.com/locate/jesp

Does exposure to erotica reduce attraction and love for romantic partners in men? Independent replications of Kenrick, Gutierrez, and Goldberg (1989) study 2[☆]

Rhonda N. Balzarini^{*}, Kiersten Dobson¹, Kristi Chin, Lorne Campbell

Department of Psychology, University of Western Ontario, London, ON, N6A 5C2, Canada

HIGHLIGHTS

- Three preregistered, high-powered replications of Kenrick et al. (1989)
- Exposed men and women in committed relationships to opposite sex erotica
- After exposure assessed ratings of attractiveness and love for partner
- Effects of original and replication studies were meta-analyzed
- Across the three studies we did not find support for the original finding.

ARTICLE INFO

Article history:

Received 8 July 2016

Revised 3 November 2016

Accepted 12 November 2016

Available online xxxxx

Keywords:

Centerfolds

Erotica

Partner attractiveness

Love

Replication

Reproducibility

ABSTRACT

Kenrick, Gutierrez, and Goldberg (1989; Study 2) demonstrated that men, but not women, in committed relationships exposed to erotic images of opposite-sex others reported lower ratings for their partner's sexual attractiveness ($d = 0.91$) and less love for their partner ($d = 0.69$) than men exposed to images of abstract art. This research has implications for understanding the possible effects of erotica on men in relationships, but has not been replicated. We conducted three preregistered, high-powered close replications, and meta-analyzed the effects of the original and replication studies. We did not find support for the original finding that exposure to attractive images of opposite-sex others affects males' ratings of their partners' sexual attractiveness or love for their partner.

© 2016 Published by Elsevier Inc.

Research by Kenrick et al. (1989) demonstrated that males exposed to attractive images of nude women (i.e., Playboy and Penthouse centerfolds) reported that their female romantic partner was less sexually attractive and reported less love for their romantic partner than those exposed to abstract art (a contrast effect). Females exposed to attractive images of nude men (i.e., Playgirl centerfolds), however, did not report that their male romantic partner was less sexually attractive or indicate less love for their male romantic partner compared to women exposed to abstract art. Guided by an evolutionary approach to human mating suggesting that physical appeal is relatively more important in mates for men

than women, Kenrick et al. (1989) concluded that men, but not women, compare the physically attractive centerfolds to their current partner and potentially view these women as possible alternative partners. An outcome of this comparison process for men is that they reduce their ratings of sexual attractiveness and feelings of love for their partner.

Since its publication, Kenrick et al.'s (1989) paper has been cited 249 times on Google Scholar, over 100 times on PsycINFO, and is a part of a growing body of work suggesting that ecological factors (e.g., social and physical ecology) have direct and immediate effects on psychological judgements about relationships. The finding that men exposed to female centerfolds show more of a contrast effect than women exposed to male centerfolds suggests a strong disadvantage of exposure to erotica for males and their romantic commitments (though see also Amelang & Pielke, 1992). The effects reported by Kenrick et al. (1989) were fairly large, with effect sizes expressed as Cohen's d equalling 0.91 for ratings of their partner's sexual attractiveness and 0.69 for participants' reported love for their partner. Given the relatively small

[☆] This research was supported by a grant awarded to Lorne Campbell by the Social Sciences and Humanities Research Council of Canada [grant number 122848].

^{*} Corresponding author.

E-mail addresses: rbalzari@uwo.ca (R.N. Balzarini), kdobson8@uwo.ca (K. Dobson), kchin25@uwo.ca (K. Chin), lcampb23@uwo.ca (L. Campbell).

¹ The first two authors contributed equally to this work.

number of men in the study (30 in total) the 95% confidence intervals of these effects are very wide: 0.16–1.66 for ratings of partner's sexual attractiveness and -0.04 – 1.43^2 for reported love for the partner. The reported effects can therefore be potentially very large or fairly small.

The purpose of the current research was to conduct independent (close) replications of Kenrick et al.'s (1989; Study 2) findings to provide additional estimates of how strongly exposure to erotica affects males' ratings of their partner's sexual attractiveness and their love for their partner. For this purpose, three high-powered (estimated power of 0.95) and pre-registered independent replications were conducted using approximately the same manipulations and measures as the original study. We anticipated that the results of the original study would replicate, such that males who were exposed to female nude centerfolds would report that their romantic partner is less attractive, and report less love for their partner, than if they were exposed to abstract art. Also consistent with the original study, we anticipated no significant effects of the type of stimuli presented on females' ratings of their partner's attractiveness and their love for their partner. As per Simmons, Nelson, and Simonsohn (2012), for all three of our replication attempts "we report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study" (p. 1).

1. Studies 1 and 2

1.1. Methods

1.1.1. Power analysis

A power analysis indicated that a sample size of 210 would be needed to find a statistically significant interaction in a 2 (gender) \times 2 (condition) analysis of variance (ANOVA) assuming a medium effect size ($f = 0.25$)³ with a power level of 0.95 (power estimated using G-Power 3.1; Erdfelder, Faul, & Buchner, 1996; Faul, Erdfelder, Lang, & Buchner, 2007). This is consistent with a developing norm for replication research that suggests obtaining a sample that is at least 2.5 times the size of the original sample to have adequate power to detect the originally reported effect (Simonsohn, 2015). Kenrick et al.'s (1989) sample size was 65 participants, indicating at least 163 participants were required for each replication study. Thus the sample size was selected to be 210, and critically, we did not conduct any analyses until we achieved at least 210 complete data points.

1.1.2. Sampling

This research was conducted in accordance with the ethical guidelines of the American Psychological Association. Unlike the original study which was conducted in the lab, participants were recruited from Amazon's Mechanical Turk (MTurk), an online crowdsourcing platform that is commonly used for psychological research. Two advertisements (one for males, one for females) were placed on the MTurk website for all MTurk users with active accounts to see, with information about the inclusion criteria (see results) as well as a link to the survey. Eligible and interested participants followed the link that redirected them to a survey hosted on Qualtrics⁴, where the letter of information and consent was presented. Informed consent was received from each participant digitally (each participant indicated they read the consent form and agreed to take part before proceeding).

² The original manuscript does not report SDs. We therefore used an approximate estimate of the SDs for the original study by calculating the SDs from the replication attempts (replication 1–3 SDs averaged per condition per group). The pooled SDs were used to approximate the confidence intervals around the effect size estimates.

³ The authors estimated the power using a smaller effect size than the original manuscript reported.

⁴ The Qualtrics QSF file and experimental stimulus available upon request at the following link: <https://osf.io/h7nxf/>

1.1.3. Procedure

The materials and procedure of these studies were conducted in accordance with the input of the first author of the original article, Dr. Douglas Kenrick. Consistent with the original study, participants were informed that the study attempts to address a controversy about whether particular works of art, photography, or cinema are artistically valuable or just offensive to "good taste." We explained that we are studying which characteristics separate "aesthetically pleasing from boring or unpleasant works" and that subjects will judge "materials which have aroused controversy with regard to their aesthetic significance." Since the materials might include nude photographs, subjects were advised to complete the study in a private place, and were reminded that they may withdraw from the study by exiting the screen at any time. The experimental males were exposed to 16 female nude centerfold images from *Playboy* and *Penthouse*, while experimental females were exposed to 16 male nude centerfold images from *Playgirl*. Control subjects viewed 16 abstract art images (e.g., Josef Albers, "Homage to the square", Jackson Pollocks, "Convergence"). Participants first indicated their sex in the demographic questionnaire, and based on this response and their random assignment to either the control or experimental condition, they either saw nude centerfolds of people of the opposite-sex, or images of abstract art. In all conditions, each image was shown for 15 s.

After viewing each image, participants were asked to rate how aesthetically pleasing the image was. After all images were shown and the aesthetic judgments were made, participants were told that there is some controversy about how relationships influence responses to art. They were told that some psychologists believe that being in a stable relationship enhances people's appreciation of art, while others feel that the deep involvement interferes with aesthetic appreciation, and still others believe that it depends on the type of relationship. Participants were then asked to respond to a questionnaire regarding how they rate their relationship. In addition, they were asked to complete a questionnaire assessing their love for their partner. Participants then filled out a suspicion probe, were fully debriefed regarding the true purpose of the study, and were provided a code to claim compensation (\$0.50).

1.2. Measures

1.2.1. Demographics

Prior to viewing the stimuli, participants were asked a series of demographic questions. Respondents were asked about their gender, sexual orientation, English fluency, age, and information about their relationship (e.g., relationship status, relationship length, whether the participant lives with their partner).⁵ Because the demographics were not reported in the original study, this questionnaire was developed by the current researchers.

1.2.2. Pleasantness

After viewing each image, three-items were used to assess how aesthetically pleasing the images were on 7-point Likert scales. Participants were asked to rate whether the images were considered artistic (1 = *highly artistic*, 7 = *not at all artistic*), pleasant (1 = *unpleasant*, 7 = *pleasant*), and socially valuable (1 = *not at all socially valuable*, 7 = *socially valuable*). Only the ratings for pleasantness were used in the analyses (pleasantness of male centerfolds $\alpha = 0.96$; pleasantness of female centerfolds $\alpha = 0.96$; pleasantness of abstract art $\alpha = 0.90$).⁶

1.2.3. Partner attractiveness

Participants were asked to answer six questions about their current romantic partner on 9-point Likert scales. Three of the questions

⁵ For a complete listing of the measures, please view the replication protocol at the following link: <https://osf.io/nrkej/>

⁶ Unless specified, procedures and measures were consistent with the original study (Kenrick et al., 1989).

included in this questionnaire were filler items (e.g., “Please rate the length of your relationship”) and three of the items assessed the partner’s attractiveness. More specifically, participants were asked to rate their partner’s attractiveness (1 = *sexually unattractive*, 9 = *sexually attractive*), to rate how attractive this person is to other males (females; 1 = *not at all*, 9 = *very much*), and to rate how desirable their partner is to other males (females) as a date (1 = *not at all*, 9 = *very much*; male rating $\alpha = 0.88$; female rating $\alpha = 0.91$).

1.2.4. Love for partner

Participants answered thirteen questions regarding their attitudes towards their partner on a 9-point Likert scale. The items used were based on Rubin’s (1970) Love Scale (e.g. “If __ were feeling badly, my first duty would be to cheer him (her) up”; 1 = *not at all true/disagree completely*, 9 = *definitely true/agree completely*; $\alpha = 0.91$).

1.2.5. Suspicion probe

A three-item suspicion questionnaire was included to assess participants’ suspicion regarding the intention of this study. Items included two open-ended questions, “What do you think the purpose of this research is?”, “At what point was this made apparent to you?”, and one item was rated on a 9-point Likert scale, “Do you believe your responses have been affected by your perception of the study purpose?” (1 = *disagree completely*, 9 = *agree completely*). The original study reported a suspicion probe but did not report any example items, thus the questionnaire in our study was compiled by the current researchers, and may not reflect the exact wording used in the original study.

1.2.6. Attention check

Two items were used to assess attention. The first item was randomly presented in the Partner Attractiveness questionnaire and asked participants to select option three from a 9-point Likert scale (1 = *not at all true/disagree completely*, 9 = *definitely true/agree completely*), and the second item asked participants, “Knowing the amount of attention you dedicated to this study, would you recommend to us that we include your data in our analyses?”, with the option to answer “yes” or “no”. The two attention check questions were not included in the original study, though such checks are typical when using an MTurk sample.

1.2.7. Analytic strategy

The series of analyses we conducted allowed us to determine (1) if the participants rated the erotic stimuli as more pleasant than the abstract art images, (2) whether ratings of attractiveness were influenced by the type of exposure (erotic stimuli vs. control), and (3) whether ratings for love were influenced by the type of exposure (erotic stimuli vs. control). The analyses followed precisely Kenrick et al.’s (1989) analytic strategy. That is, for each participant the pleasantness ratings for the stimuli were averaged to create a pleasantness score. A 2×2 ANOVA was then used to examine whether there were any differences between genders and conditions on the ratings of pleasantness of the stimuli. Next, the three partner attractiveness ratings (excluding filler items) were aggregated to form a partner attractiveness score, which was used as the dependent variable in another 2×2 ANOVA examining the differences between genders and conditions on ratings of partner attractiveness. Lastly, we aggregated the 13 items from Rubin’s (1970) Love Scale to create a love for partner score which served as the dependent variable in a final 2×2 ANOVA on the differences between genders and conditions on love for their partner. To be conservative, all simple main effects associated with these ANOVAs utilized Bonferroni adjustments.

2. Study 1 results

2.1. Demographics

A total of 438 participants accessed the online study. Of those, 316 consented to participation. Only 241 of those people met the inclusion criteria (2 removed for not speaking English fluently, 38 for indicating they are single, and 35 for indicating they do not live with their partner). In addition, nine participants were removed for failing the attention check questions, five for indicating they are lesbian or gay (which may alter their responses to the centerfold conditions),⁷ three participants were removed for careless responding, and one for indicating extreme suspicion of the true nature of the study in the suspicion check. This left a total sample of 223 (124 male, 99 female) participants. The average age of the sample was 35.43 years ($SD = 11.10$ years), and the average relationship length was 9.09 years ($SD = 9.41$ years).⁸

2.2. Pleasantness

A 2 (gender: male, female) \times 2 (art condition: abstract art, nude centerfolds) ANOVA was run for participants’ average ratings of the pleasantness of the images. A main effect of gender was found ($F(1, 219) = 10.19, p = 0.002; \eta^2 = 0.044$), such that males ($M = 5.07, SE = 0.11$) rated photos as more pleasant than females ($M = 4.55, SE = 0.12$). However, the main effect of art condition was not significant ($F(1, 219) = 0.121, p = 0.728; \eta^2 = 0.001$). An interaction of gender and art condition also emerged ($F(1, 219) = 22.29, p < 0.001; \eta^2 = 0.092$). An analysis of the simple main effects revealed that males rated the nude centerfolds as significantly more pleasant than the abstract art ($F(1, 219) = 14.71, p < 0.001; \eta^2 = 0.063$), while females rated the abstract art as significantly more pleasant than the nude centerfolds ($F(1, 219) = 8.49, p = 0.004; \eta^2 = 0.037$), and males rated the nude centerfolds as more pleasant than females did ($F(1, 219) = 28.18, p < 0.001; \eta^2 = 0.114$). This is partially consistent with the original study, in which both males and females rated the nude centerfolds as more pleasant than the abstract art. See Table 1 for the means and standard deviations in each condition.

2.3. Partner attractiveness

A 2 (gender: male, female) \times 2 (art condition: abstract art, nude centerfolds) ANOVA was run for participants’ ratings of their partner’s attractiveness. No significant main effects (gender: $F(1, 219) = 0.06, p = 0.805; \eta^2 < 0.001$; art condition: $F(1, 219) = 0.09, p = 0.771; \eta^2 < 0.001$) or an interaction ($F(1, 219) = 0.03, p = 0.872; \eta^2 < 0.001$) were found. Analyses of the simple main effects did not reveal any significant differences between groups. This is inconsistent with the original study, which found a difference between males in the abstract art and nude centerfold conditions. See Table 2 for the means and standard deviations in each condition.

2.4. Love for partner

A 2 (gender: male, female) \times 2 (art condition: abstract art, nude centerfolds) ANOVA was run for participants’ ratings of their love for their partner. No significant main effects (gender: $F(1, 219) = 2.37, p = 0.125; \eta^2 = 0.011$; art condition: $F(1, 219) = 0.12, p = 0.730; \eta^2 = 0.001$) or an interaction ($F(1, 219) = 1.19, p = 0.276; \eta^2 = 0.005$) were found. However, an analysis of the simple main effects revealed that in the abstract art condition, females reported more love for their

⁷ The analyses for all studies were run with and without bisexual participants. Results remained the same with these participants included, unless otherwise specified.

⁸ A copy of the data files, syntax, and outputs for all 3 replication studies are available at <https://osf.io/b47jp/>.

Table 1

Descriptive information regarding average pleasantness ratings across conditions in the original study and the three replication attempts.

Study	Male			Female		
	Abstract art	Nude centerfold	Non-nude photographs	Abstract art	Nude centerfold	Non-nude photographs
Kenrick et al. (1989)	4.0 (0.88) <i>n</i> = 15	4.9 (1.21) <i>n</i> = 15	–	3.9 (1.02) <i>n</i> = 17	4.9 (1.58) <i>n</i> = 18	–
Study 1	4.66 (0.93) <i>n</i> = 64	5.47 (1.20) <i>n</i> = 60	–	4.90 (1.08) <i>n</i> = 58	4.20 (1.59) <i>n</i> = 41	–
Study 2	4.59 (0.88) <i>n</i> = 85	5.67 (1.20) <i>n</i> = 85	–	4.72 (0.97) <i>n</i> = 52	4.41 (1.54) <i>n</i> = 41	–
Study 3	4.71 (0.84) <i>n</i> = 63	5.54 (1.23) <i>n</i> = 58	5.80 (0.86) <i>n</i> = 69	4.45 (1.01) <i>n</i> = 50	4.07 (1.62) <i>n</i> = 54	4.77 (1.27) <i>n</i> = 50

Note. The standard deviation is provided in parentheses. Additionally, the original manuscript reports means with only 1 decimal place and SDs were not reported. We used an approximate estimate of the SDs for the original study by calculating the SDs from the replication attempts (replication 1–3 SDs averaged per condition per group). Lastly, the original study reports 30 males and 35 females— we assumed that in the original study there were 15 males in each group, and arbitrarily assigned one of the female conditions 17, and the other 18.

Table 2

Descriptive information regarding average partner attractiveness ratings across conditions in the original study and the three replication attempts.

Study	Male			Female		
	Abstract art	Nude centerfold	Non-nude photographs	abstract art	Nude centerfold	Non-nude photographs
Kenrick et al. (1989)	26.2 (4.40) <i>n</i> = 15	22.1 (4.38) <i>n</i> = 15	–	22.5 (4.64) <i>n</i> = 17	22.2 (4.12) <i>n</i> = 18	–
Study 1	21.52 (4.19) <i>n</i> = 64	21.23 (4.46) <i>n</i> = 60	–	21.57 (5.09) <i>n</i> = 58	21.49 (4.62) <i>n</i> = 41	–
Study 2	21.35 (5.11) <i>n</i> = 85	21.82 (4.42) <i>n</i> = 85	–	22.00 (3.95) <i>n</i> = 52	20.61 (4.34) <i>n</i> = 41	–
Study 3	20.86 (3.90) <i>n</i> = 63	21.47 (4.27) <i>n</i> = 58	21.93 (3.90) <i>n</i> = 69	19.46 (4.89) <i>n</i> = 50	21.44 (3.40) <i>n</i> = 54	21.84 (3.67) <i>n</i> = 50

Note. The standard deviation is provided in parentheses.

partner than males ($F(1, 219) = 3.90, p = 0.050; \eta^2 = 0.017$).⁹ No significant difference was found between males and females in the nude centerfold condition ($F(1, 219) = 0.09, p = 0.764; \eta^2 < 0.001$). No other significant differences were found between groups. This is inconsistent with the results of the original study, which found a difference between males in the abstract art and nude centerfold conditions. See Table 3 for the means and standard deviations in each condition.

3. Study 2 results

3.1. Demographics

There were 484 participants who opened the online study. Of those, 423 consented to participation. Only 338 of those people met the inclusion criteria (14 removed for not speaking English fluently, 25 for indicating they are single, and 46 for indicating they do not live with their partner). In addition, 51 participants were removed for failing the attention check questions, eight for indicating they are lesbian or gay (which may alter their responses to the centerfold conditions), eight participants were removed for careless responding, seven for indicating that their relationship length was <3 months or >100 years, and one for indicating their age was <18 years old. This left a total sample of 263 (170 male, 93 female) participants. The average age of the sample was 35.67 years ($SD = 10.58$ years), and the average relationship length was 8.81 years ($SD = 8.68$ years).

3.2. Pleasantness

A 2 (gender: male, female) \times 2 (art condition: abstract art, nude centerfolds) ANOVA was run for participants' average ratings of the pleasantness of the images. Significant main effects of gender ($F(1, 259) = 14.95, p < 0.001; \eta^2 = 0.055$) and art condition ($F(1, 259) = 6.82, p = 0.010; \eta^2 = 0.026$) were found, such that males ($M = 5.13, SE =$

0.09) rated photos as more pleasant than females ($M = 4.57, SE = 0.12$), and nude centerfolds ($M = 5.04, SE = 0.11$) were rated as more pleasant than abstract art ($M = 4.66, SE = 0.10$). In addition, an interaction of gender and art condition ($F(1, 259) = 22.58, p < 0.001; \eta^2 = 0.080$) emerged. Consistent with Study 1, an analysis of the simple main effects revealed that males viewed the nude centerfolds as significantly more pleasant than the abstract art ($F(1, 259) = 38.68; p < 0.001, \eta^2 = 0.130$), and males viewed the nude centerfolds as significantly more pleasant than females did ($F(1, 259) = 34.49; p < 0.001, \eta^2 = 0.118$). There was no significant difference between females' pleasantness ratings in the abstract art and nude centerfold conditions ($F(1, 259) = 1.76, p = 0.185; \eta^2 = 0.007$), although these results trended in the same direction as the previous study. See Table 1 for the means and standard deviations in each condition.

3.3. Partner attractiveness

A 2 (gender: male, female) \times 2 (art condition: abstract art, nude centerfolds) ANOVA was run for participants' ratings of their partner's attractiveness. No significant main effects (gender: $F(1, 259) = 0.23, p = 0.632; \eta^2 = 0.001$; art condition: $F(1, 259) = 0.61, p = 0.437; \eta^2 = 0.002$) or an interaction ($F(1, 259) = 2.48, p = 0.116; \eta^2 = 0.009$) were found. Analyses of the simple main effects did not reveal any significant differences between groups. These results are consistent with Study 1. See Table 2 for the means and standard deviations in each condition.

3.4. Love for partner

A 2 (gender: male, female) \times 2 (art condition: abstract art, nude centerfolds) ANOVA was run for participants' ratings of their love for their partner. No significant main effects (gender: $F(1, 259) = 0.51, p = 0.475; \eta^2 = 0.002$; art condition: $F(1, 259) = 0.15, p = 0.699; \eta^2 = 0.001$) or an interaction ($F(1, 259) = 1.40, p = 0.238; \eta^2 = 0.005$) were found. An analysis of the simple main effects revealed that in the abstract art condition, females did not report significantly more love

⁹ This effect was nonsignificant when bisexual participants were excluded ($t(201) = 1.37, p = 0.173; d = 0.19$).

Table 3

Descriptive information regarding average love for partner ratings across conditions in the original study and the three replication attempts.

Study	Male			Female		
	Abstract art	Nude centerfold	Non-nude photographs	Abstract art	Nude centerfold	Non-nude photographs
Kenrick et al. (1989)	98.8 (15.99) <i>n</i> = 15	87.2 (16.51) <i>n</i> = 15	–	85.0 (16.75) <i>n</i> = 17	82.3 (16.32) <i>n</i> = 18	–
Study 1	89.19 (16.21) <i>n</i> = 64	92.48 (17.75) <i>n</i> = 60	–	95.22 (15.79) <i>n</i> = 58	93.51 (17.98) <i>n</i> = 41	–
Study 2	90.94 (18.12) <i>n</i> = 85	92.66 (14.81) <i>n</i> = 85	–	95.04 (16.95) <i>n</i> = 52	91.65 (16.25) <i>n</i> = 40	–
Study 3	92.52 (13.65) <i>n</i> = 63	90.07 (16.96) <i>n</i> = 58	93.68 (17.60) <i>n</i> = 69	90.60 (17.51) <i>n</i> = 50	94.19 (14.73) <i>n</i> = 54	93.78 (16.02) <i>n</i> = 50

Note. The standard deviation is provided in parentheses.

for their partner than males ($F(1, 259) = 1.97, p = 0.162; \eta^2 = 0.008$), although results trended in the same direction as the previous study. No other significant differences were found between groups. See Table 3 for the means and standard deviations in each condition.

4. Study 3

4.1. Methods

4.1.1. Power analysis

In this study we added a third art condition as participants (particularly females) in the first two studies did not rate nude centerfolds as more pleasant than abstract art, which is inconsistent with the original study. Therefore, the third art condition we added included attractive non-nude opposite-sex images. This analysis involved a 2×3 ANOVA to examine the main effects and interaction between gender and condition. A power analysis indicated that a sample size of 251 would be needed to find a statistically significant interaction assuming a medium effect size ($f = 0.25$) with a power level of 0.95 (power estimated using G-Power 3.1; Erdfelder et al., 1996; Faul et al., 2007).

4.1.2. Sampling and procedure

The sampling method and procedure used was the same as Studies 1 and 2, with the exception of the addition of the third art condition. These images were gathered from publicly available websites, including Google image searches and Pinterest, and included photographs used in calendars, advertisements, and boudoir photos. Sixteen photographs of non-nude males and 16 images of non-nude females (clothed in such a way that both breasts and genitalia were covered by clothing) were selected based on the quality of the image, the attractiveness of the person in the photo, and how artistic the photo appeared to be (to fit with the cover story), based on the perspective of seven raters in informal discussions. Participants who were randomly sorted into this condition viewed each of the 16 photos for 15 s each, and provided ratings of each photo, followed by answering the remaining questionnaires.

4.1.3. Measures

The measures used were the same as those used in studies 1 and 2, with the addition of a question asking participants if they had participated in similar study to this in the past. As this was our third replication attempt with an MTurk sample, we wanted to ensure that there were no participants taking part in multiple replication attempts.

4.1.4. Analytic strategy

The analytic strategy used was the same as studies 1 and 2, with the exception of an additional condition (e.g., non-nude photographs). Thus, three 2×3 ANOVAs (gender: male, female; art condition: abstract art, nude centerfolds, non-nude photographs) were run, to examine the same effects with the third art condition included.

4.2. Results

4.2.1. Demographics

There were 640 participants who opened the online study. Of those, 568 consented to participation. Only 433 of those people met the inclusion criteria (28 removed for not speaking English fluently, 51 for indicating they are single, and 56 for indicating they do not live with their partner). In addition, 68 participants were removed for failing the attention check questions, nine for indicating they are lesbian or gay, two participants were removed for careless responding, nine for indicating they had participated in a similar study in the past, and one for indicating their age was <18 or over 100. This left a total sample of 344 (190 male, 154 female) participants. The average age of the sample was 35.26 years ($SD = 11.06$ years), and the average relationship length was 9.14 years ($SD = 9.71$ years).

4.2.2. Pleasantness

A 2 (gender: male, female) \times 3 (art condition: abstract art, nude centerfolds, non-nude photographs) ANOVA was run for participants' average ratings of the pleasantness of the images. Main effects of gender ($F(1, 338) = 53.57, p < 0.001; \eta^2 = 0.137$), and art condition ($F(2, 338) = 11.16, p < 0.001; \eta^2 = 0.062$) were found, such that males ($M = 5.35, SE = 0.08$) rated images as more pleasant than females ($M = 4.43, SE = 0.09$), and non-nude photographs were rated as more pleasant ($M = 5.28, SE = 0.11$) than abstract art ($M = 4.58, SE = 0.11, t(338) = 4.61, p < 0.001; d = 0.50$) and nude centerfolds ($M = 4.80, SE = 0.11, t(338) = 3.16, p = 0.005; d = 0.42$). In addition, an interaction of gender and art condition ($F(2, 338) = 7.82, p < 0.001; \eta^2 = 0.044$) was found. An analysis of the simple main effects revealed that males rated nude centerfolds as significantly more pleasant than females ($F(1, 338) = 45.27, p < 0.001; \eta^2 = 0.118$), and rated non-nude photographs as more pleasant than females ($F(1, 338) = 22.77, p < 0.001; \eta^2 = 0.063$). Males rated nude centerfolds as significantly more pleasant than abstract art ($t(338) = 3.94, p < 0.001; d = 0.72$), and rated non-nude photographs as more pleasant than abstract art ($t(338) = 5.41, p < 0.001; d = 0.95$). Females rated the non-nude photographs as more pleasant than the nude centerfolds ($t(338) = 3.11, p = 0.006; d = 0.61$). See Table 1 for the means and standard deviations for each condition.

4.2.3. Partner attractiveness

A 2 (gender: male, female) \times 3 (art condition: abstract art, nude centerfolds, non-nude photographs) ANOVA was run for participants' ratings of their partner's attractiveness. A main effect of art condition was found ($F(2, 338) = 5.64, p = 0.004; \eta^2 = 0.032$), such that participants who viewed nude centerfolds ($M = 21.46, SE = 0.38$) or non-nude photographs ($M = 21.88, SE = 0.37$) rated their partners as more attractive than those who saw abstract art ($M = 20.16, SE = 0.38, t(338) = 2.41, p = 0.049; d = 0.32^{10}; t(338) = 3.24, p = 0.004$;

¹⁰ This effect was marginally significant when bisexual participants were excluded from analyses ($t(314) = 2.14, p = 0.10; d = 0.29$).

$d = 0.43$, respectively). No significant main effect of gender was found ($F(1, 338) = 1.32, p = 0.251; \eta^2 = 0.004$) or interaction of gender by art condition ($F(2, 338) = 1.05, p = 0.353; \eta^2 = 0.006$). An analysis of the simple main effects revealed that females who saw nude centerfolds or non-nude photographs rated their partners as more attractive than females who saw abstract art ($t(338) = 2.51, p = 0.037^{11}; d = 0.49; t(338) = 2.96, p = 0.010; d = 0.59$, respectively). See Table 2 for the means and standard deviations for each condition.

4.2.4. Love for partner

A 2 (gender: male, female) \times 3 (art condition: abstract art, nude centerfolds, non-nude photographs) ANOVA was run for participants' ratings of their love for their partner. No significant main effects (gender: $F(1, 338) = 0.19, p = 0.663; \eta^2 = 0.001$; art condition: $F(2, 338) = 0.56, p = 0.575; \eta^2 = 0.003$) or interaction ($F(2, 338) = 1.01, p = 0.364; \eta^2 = 0.006$) were found. An analysis of the simple main effects revealed no significant differences between groups. See Table 3 for the means and standard deviations for each condition.

5. Exploratory analyses

The original study recruited students who were enrolled at a large state university or community college, and who were married or involved in a similar live-in situation. All three of our replication attempts consisted of MTurk workers, whose average age was 35, who were also married or involved in similar live-in relationship with an average relationship length of 9 years. Although no information on the average age or relationship length of participants in the original study were provided, differences in recruitment methods most likely led to the replication samples being older and thus having been in a relationship for a longer duration of time than participants in the original study. We conducted a series of exploratory analyses assessing whether these potential differences in age or relationship length were predictors of ratings of pleasantness, partner attractiveness, and love for partner, along with art condition and gender, plus the interactions for each term. We analyzed the data using both age, age centered (centered around age 20, a proxy for the average age of college students, per reviewer request), and relationship length separately as predictor variables and conducted the analyses across each study of the replication data, and for the merged dataset of all three replication attempts to provide more power.

5.1. Age

In Study 1, a main effect of age was found for love for partner ($b = 0.377, p = 0.049$), indicating that as age increased, love increased. Please note that results for age and age centered did not differ for these analyses, so we will discuss age generally from here on. No other significant effects or interactions for age emerged in Study 1. In Study 2, a main effect of age was found for partner attractiveness ratings ($b = -0.127, p = 0.005$), indicating that as age increased, partner attractiveness ratings decreased. A main effect for age was also found for love for partner ($b = -0.344, p = 0.038$), though in Study 2, as age increased, love for partner decreased, which is contrary to Study 1 results. No significant effects emerged for age for Study 3 or in the merged dataset. These exploratory analyses suggest that the effects for age across the studies are inconsistent, and demonstrate that no simple age effect can account for the differences in results between the replication attempts and the original study.

5.2. Relationship length

In Study 1, an interaction of relationship length and condition emerged for pleasantness ratings ($b = 0.057, p = 0.014$), indicating

that relationship length had a larger effect for individuals in the center-fold condition. Additionally, a marginally significant interaction between relationship length, condition, and gender emerged for pleasantness ratings ($b = -0.072, p = 0.054$), such that males rated the nude centerfolds as more pleasant than the abstract art, and this was particularly true for males who had been in a relationship longer than males who have been in a relationship for a shorter period of time. A main effect for relationship length was also found for love for partner ($b = 0.452, p = 0.012$), indicating that as relationship length increased, love increased. No significant effects emerged for relationship length in Study 2 or Study 3. Lastly, in the merged data, an interaction of relationship length, condition, and gender emerged for pleasantness ratings ($b = -0.044, p = 0.039$), such that males rated the nude centerfolds as more pleasant than the abstract art, and this was particularly true for males who have been in a relationship for longer period of time. As this interaction did not emerge in Study 2 or 3, it appears to derive from a strong effect which emerged in Study 1 and thus was inconsistent across the studies.¹²

6. Meta-analysis of combined replication samples

To obtain a more precise estimate of the effect size of the primary comparison in the Kenrick et al. (1989) Study 2 (i.e., between men's attraction and love in the nude centerfold versus abstract art conditions), a random effects meta-analysis was conducted for each dependent variable using the metafor package in R (Viechtbauer, 2010; see Figs. 1 and 2). Each meta-analysis contained four independent effects (i.e., $k = 4$) with a total of 218 participants in the nude centerfold condition and 227 participants in the abstract art condition. For the meta-analysis on partner attractiveness ratings there was a weighted mean effect size of $d = 0.02$, 95% CI $[-0.21, 0.24]$, and marginally significant variation in the effect sizes between studies was evident (i.e., the magnitude of the effect sizes differed somewhat between studies), $Q(3) = 6.82, p = 0.078$. For the meta-analysis on love for partner ratings there was a weighted mean effect size of $d = 0.02$, 95% CI $[-0.22, 0.26]$, and there was no evidence of significant variation in the effect sizes between studies, $Q(3) = 5.77, p = 0.124$. The results of the meta-analyses therefore suggest the effect sizes of the main effects of interest are very close to, and not significantly different from, zero. Furthermore, the confidence intervals are consistent with a small positive effect all the way to a small negative effect for partner attractiveness ratings, and love for partner ratings.¹³

7. Discussion

The results of these three close replications of Study 2 by Kenrick et al. (1989) are not consistent with the original results. Specifically, in the original study, both men and women found the erotic images to be more pleasant than abstract art. In our replication attempts, men consistently rated nude centerfolds as more pleasant than abstract art, though females rated nude centerfolds as less pleasant than abstract art (though this difference was not significant in Study 2 and 3). In Study 3 a third condition was added (e.g., non-nude, yet attractive males/females), and both men and women found the non-nude photographs more pleasant than the nude centerfolds or abstract art. More importantly, in the original study men who were exposed to nude centerfolds rated their partners as less attractive, and reported less love for their partner, than those who were exposed to abstract art, while exposure to nude centerfolds versus abstract art had no effect on women. In Study 1 and 2 we found no effect of exposure to erotica on ratings of partner's attractiveness, and this was true for both men and women. However, in Study 3 we found that individuals exposed to images of nude centerfolds or of

¹¹ This effect was marginally significant when bisexual participants were excluded from analyses ($t(314) = 2.28, p = 0.069; d = 0.31$).

¹² A copy of the data files, syntax, and outputs for the exploratory analyses are available at <https://osf.io/8xcea/>

¹³ Meta-analyses for males and females available on the OSF at: <https://osf.io/kzuqa/>

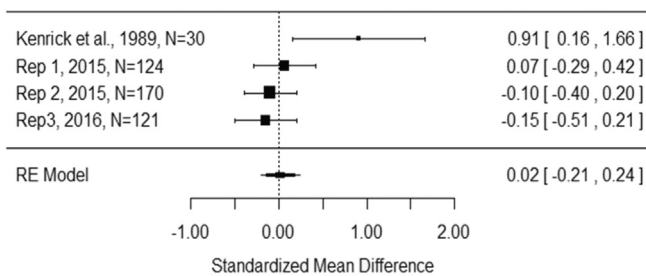


Fig. 1. Random effects meta-analysis on males' partner attractiveness ratings comparing participants in the abstract art versus nude centerfold condition. Error bars are 95% confidence intervals.

non-nude individuals actually rated their partner as *more* attractive (result was not significant for males). Interestingly, across the three replication attempts, we found no effect of exposure to erotica on ratings of love for one's partner. Thus exposure to nude centerfolds *may not* influence attraction to and love for one's partner, for males or females.

The failure to replicate Kenrick et al.'s (1989) Study 2 results warrants a discussion of the possible explanations for the observed conflicting findings. First, it is necessary to point out that the original study was published in 1989. At the time, exposure to sexual content may not have been as available (Reichert, Lambiase, Morgan, Carstarphen, & Zavoina, 1999), and was less common among women, who may have been led to believe that pornography exploits, objectifies, and degrades women (Kerner, 2011). Today, exposure to nude images is relatively more pervasive and women are a growing body of consumers of pornography (Berman, 2014). Thus, the results for the current replication studies may differ due to differences in exposure, access, and acceptance of erotica then versus now. To assess whether these factors can account for the difference in our findings from the original study, future researchers should consider including them as potential moderators.

It is also possible that methodological differences between the original study and the replication attempts may account for differences in the results. The original study had a university sample, and participants viewed images on a projection screen and completed their ratings with pen-and-paper questionnaires in small groups in a lab. In the current replications, the participants were recruited online through MTurk, and involvement consisted of completing the study individually via an online survey provider. In addition, the stimuli (images) presented to the participants were not the same images that were presented in the original study. After consulting with the original author, it was decided that updated nude centerfolds would be used to be consistent with the experience of the original study rather than the exact materials. Although we believe it to be improbable, it is possible that these methodological differences contributed to our inability to replicate the findings of the original study. If such were the case, then the original results either depend crucially on the sample (e.g., college vs. community sample), the modality of taking the survey (e.g., in lab vs. online; with others vs. alone), changes in imagery over time, or potentially by a

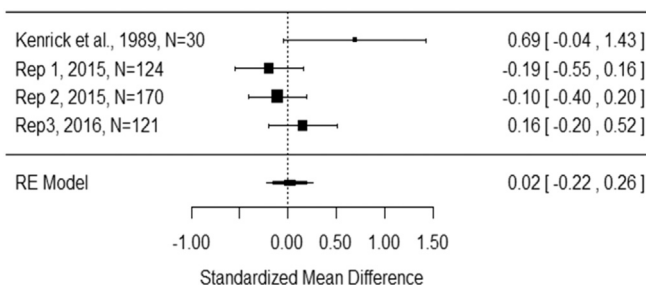


Fig. 2. Random effects meta-analysis on males' love for partner ratings comparing participants in the abstract art versus nude centerfold condition. Error bars are 95% confidence intervals.

suppressor variable, such as individuals' tendencies towards erotophobia/erotophilia.

To assess whether some of these methodological differences can account for the difference in our findings, future researchers should consider these possible moderators before starting the replication study (see Brandt et al., 2014) and should include additional measures that can assist in explaining why the results changed (e.g., Sexual Opinion Survey; Fisher, White, Byrne, & Kelley, 1988). With that said, if any of these alternatives are true, then it is unclear whether the effect is robust enough to have an impact outside of the laboratory (e.g., in people's actual relationships), or on people generally (e.g., if in future research the effect emerges only in erotophobic individuals).

Finally, one possible explanation for our inability to replicate the effects found in the original study is because the effects are either very small or simply do not exist. Discussed previously, it is possible that the small sample size of the original study led to an artificial inflation of the effect size. Our meta-analyses indicate that the confidence interval for the effect size across the original study and all three replication attempts includes zero for the effects of condition on males' ratings of their partner's attractiveness (95% CI: -0.21, 0.24) and their love for their partner (95% CI: -0.22, 0.26). More specifically, the overall effect sizes expressed as Cohen's *d* were 0.02 for participants' reported attraction to their partner, and 0.02 for their love for their partner.

What can be concluded about the role of exposure to nude centerfolds on male's ratings of attractiveness and love for their partner? Based on these three preregistered, high-powered close replications and reviews of the meta-analyzed effects of the original and replication studies, we were unable to find support for the original finding that exposure to attractive images of opposite-sex others affects males' ratings of their partners' sexual attractiveness or love for their partner. More specifically, the best current effect size estimate is considerably smaller than envisioned in the original study. This does not mean that the effect is not there or was not there when the original authors conducted the study, though it could suggest that the earlier reported contrast effect in men is not as robust as originally thought.

References

- Amelang, M., & Pielke, M. (1992). Effects of erotica upon men's and women's loving and liking responses for their partners. *Psychological Reports*, 71, 1235–1245.
- Berman, L. (2014, January 22). Men and women differ in their tastes for adult material [blog post]. (Retrieved from <http://www.everydayhealth.com/columns/dr-laura-berman-on-love-and-sex/men-and-women-differ-in-their-tastes-for-erotica/>)
- Brandt, M. J., Ijzerman, H., Dijksterhuis, A., Farach, F. J., Geller, J., Giner-Sorolla, R., ... Van't Veer, A. (2014). The replication recipe: What makes for a convincing replication? *Journal of Experimental Social Psychology*, 50, 217–224.
- Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments, & Computers*, 28, 1–11.
- 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, 175–191.
- Fisher, W. A., White, L. A., Byrne, D., & Kelley, K. (1988). Erotophobia-erotophilia as a dimension of personality. *Journal of Sex Research*, 25, 123–151.
- Kenrick, D. T., Gutierrez, S. E., & Goldberg, L. L. (1989). Influence of popular erotica on judgments of strangers and mates. *Journal of Experimental Social Psychology*, 25, 159–167.
- Kerner, I. (2011, April 28). Do women like porn as much as men? [blog post]. (Retrieved from <http://thechart.blogs.cnn.com/2011/04/28/do-women-like-porn-as-much-as-men/>)
- Reichert, R., Lambiase, J., Morgan, S., Carstarphen, M., & Zavoina, S. (1999). Cheesecake and beefcake: No matter how you slice it, sexual explicitness in advertising continues to increase. *Journalism and Mass Communication Quarterly*, 76, 7–20.
- Rubin, Z. (1970). Measurement of romantic love. *Journal of Personality and Social Psychology*, 16, 265–273.
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2012). A 21 Word Solution. *Dialogue: The Official Newsletter of the Society for Personality and Social Psychology*, 26, (pp. 4–7), 4–7.
- Simonsohn, U. (2015). Small telescopes: Detectability and the evaluation of replication results. *Psychological Science*, 26, 59–569.
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36, 1–48.