Supplementary Material

S1 Methodological Details

Participants

The sample was 80% Caucasian, 16.7% Asian or Asian American, and 3.3% from other ethnic backgrounds. All but six of the yoked pairs were matched on ethnicity. Results excluding the unmatched pairs did not differ substantially from those reported here. None of the participants had distinctive facial hair or markings.

Measures and Procedures

Assessing Facially-Triggered Transference using A′. We indexed the extent to which participants were likely to judge partner-similar (vs. yoked-similar) faces as possessing a trait by computing A′. Partner-similar faces were assigned as the "signal" and yoked-similar faces were assigned as the "noise." For example, partner-similar faces were counted as a "hit" if they were judged as trustworthy. Accordingly, yoked-similar faces were counted as a "false alarm" if they were judged as trustworthy. Using the hit and false alarm rates, we calculated A′ scores (Snodgrass & Corwin, 1988) for each of the six traits (i.e., trustworthy, supportive, intelligent, attractive, accepting, aggressiveness). Chance responding (i.e., no tendency to judge partner-similar (vs. yoked-similar) faces as more likely to possess a trait) is indexed by an A′ of .5. An A′ significantly (p<.05) greater than .5 indicates a tendency to evaluate partner-similar faces as more likely to possess the trait than yoked-similar faces. Finally, we averaged all six A′s (after reverse scoring $A'_{aggressiveness}$) to index *facially-triggered transference* ($A'_{aggregate}$; α=.88). An $A'_{aggregate}$ significantly (p<.05) greater than .5 indicates a tendency to evaluate partner-similar faces more positively than yoked-similar faces.

Additional Information on Data Analysis. We first paired couples (henceforth called "dyads") and formed "couple pairs." We then created yoked pairs between same-sex participants. Because data points for each participant were nested within dyads, which were further nested within couple pairs, we used linear mixed models (LMMs) to account for interdependency among data points. Dyad and couple pair were included in the model as random variables and sex was included as a fixed variable.

We ran separate LMMs for $A'_{aggregate}$ and A′ scores for each of the six traits. We obtained the mean estimates for A′ from the LMMs, therefore, statistically controlled for the nested design. We conducted one-sample *t*-tests on these estimates, comparing each to the chance level of .5 (see Table S1).

LMMs in SPSS use Satterthwaite's (1946) approximation to estimate the degrees of freedom associated with the intercept and slopes, resulting in non-integer degrees of freedom. Degrees of freedom for mean $A'_{aggregate}$'s for men and women were estimated by the model to be larger than degrees of freedom for comparing men and women because the variance component for the couple pairs was estimated to be zero and hence, dropped out of the model.

S2 Additional Analyses

Calculating Effect Sizes

There is no established method of calculating effect sizes for complex non-independent designs such as the LMMs used in the present study (Klein, 2004). However, to give the reader a general sense of the magnitude of the effects, we calculated the effect size in units of residual variation, estimated by the full model. Our estimate of effect size is equivalent to Cohen's *d*, except that we used the residual standard deviation. Specifically, we estimated the effect size using the following formulas:

For comparing a sample mean to chance: d=(Estimated mean-0.5)/SD_{res}

For comparing two sample means: d=(Estimated $mean_1$-Estimated $mean_2$)/SD_{res}

where SD_{res} is the residual standard deviation—i.e. the square root of the error variance.

Assessing Facially-Triggered Transference in the Absence of Awareness of the Resemblance

Subjective awareness. Participants reported whether the novel faces resembled anyone whom they knew, and if yes, whom the faces resembled. We counted those participants who reported that one or more of the faces reminded them of their partner as subjectively aware. We did not count those participants who mentioned multiple SOs including the partner (*N*=3)

as subjectively aware because they appeared to be guessing. Nonetheless, when these individuals were included in the analyses, the results did not differ substantially from those reported below.

Subjective methods of awareness indicated that 14 participants (9 female) expressed awareness of the resemblance between the novel faces and their partner. For these participants, facially-triggered transference, as reflected by $A'_{aggregate}$, was .61, compared to .56 for those who did not. When we entered subjective awareness as a factor in the model (0=awareness; 1=no awareness) predicting $A'_{aggregate}$, neither the main effect of subjective awareness nor its interaction with sex was statistically significant (*t*s<1). Critically, when we excluded participants who expressed subjective awareness, $A'_{aggregate}$ remained significantly above chance for women (M=.62, t(40.95)=3.23, p<.01) and at chance for men (M=.50, t<1) (see Table S2).

Objective awareness. A subset of participants (n=46) completed a measure of objective awareness that reflects the sensitivity to consciously discriminate partner-similar faces from yoked-similar faces, indexed by $A'_{awareness}$. Objective awareness was above chance for both women (M=.84; t(45)=9.28, p<.001, d=1.89) and men (M=.80; t(45)=7.99, p<.001, d=1.67), and did not significantly differ (t<1) between the genders. We included $A'_{awareness}$ and its interaction with sex in LMMs (along with dyad, couple pair, and sex). The interaction between sex and $A'_{awareness}$ was statistically significant (b=.62, t(40.17)=2.34, p<.05), which indicated that objective awareness was related to $A'_{aggregate}$ for women (b=.69, t(22)=3.49, p<.01), but not for men (t<1). $A'_{aggregate}$ was also stronger for women than for men (t(39.48)=1.96, p=.06).

Because the relation between awareness ($A'_{awareness}$) and transference effects ($A'_{aggregate}$) varied for men and women, we ran two separate LMMs, one for women and one for men, to statistically control for $A'_{awareness}$. One-sample *t*-tests were performed using the estimated means and standard errors to compare the means to chance (.5). $A'_{aggregate}$ remained significantly above chance for women (M=.61, t(22)=3.67, p=.001) and at chance for men (M=.52, t<1), after controlling for $A'_{awareness}$. We followed the data analytic techniques

described above to investigate facially-triggered transference for each individual trait while statistically controlling for objective awareness (see Table S3).

Finally, participants who expressed subjective awareness performed significantly (p<.05) better on the measure of objective awareness ($A'_{awareness}$=.92) than those who did not (M=.78). However, when participants who expressed awareness were excluded from the analyses, the results controlling for $A'_{awareness}$ did not differ substantially from those reported in Table S3.

Explaining Gender Differences

Could the gender difference in facially-triggered transference (as reflected by $A'_{aggregate}$) have arisen because hair is a more personally identifying attribute for women than men and in the present study hair was digitally removed from all photographs? Because reaction times and error rates on the objective awareness task did not differ between genders (ts<1), the observed sex difference in facially-triggered transference was unlikely the result of digitally removing the partner's hair.

References

Kline, R. B. (2004). *Beyond Significance Testing.* American Psychological Association, Washington, DC.

Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics, 2*, 110-114.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117*, 34-50.

Tables

Table S1. Tendency to judge partner-similar (vs. yoked-similar) faces as possessing a particular trait, reflected by mean A′ scores for each of the six trait judgments and their aggregate (i.e., facially-triggered transference), for women and men.

	Women's A'				Men's A'				Women's A' compared to men's	
	Mean	SE	*t*	*p*	Mean	SE	*t*	*p*	*t*	*p*
Trustworthy	0.66	0.04	4.23	0.0003	0.49	0.04	-0.27	0.7877	3.88	0.0006
Supportive	0.65	0.04	3.91	0.0003	0.55	0.04	1.19	0.2412	2.33	0.0275
Intelligent	0.64	0.04	3.70	0.0005	0.49	0.04	-0.31	0.7574	3.31	0.0026
Attractive	0.64	0.04	3.41	0.0012	0.58	0.04	1.91	0.0621	1.09	0.2854
Accepting	0.60	0.04	2.35	0.0223	0.52	0.04	0.45	0.6558	1.33	0.1893
Aggressive	0.45	0.04	-1.44	0.1551	0.52	0.04	0.62	0.5409	-1.45	0.1530
Aggregate[a]	**0.62**	**0.03**	**4.11**	**0.0001**	**0.52**	**0.03**	**0.54**	**0.5895**	**2.83**	**0.0085**

$N = 57$.

Notes. A' is a sensitivity measure adjusted for response bias. An A' of .5 reflects chance responding. An A' significantly greater than .5 reflects the tendency to judge partner-similar (vs. yoked-similar) faces as possessing the trait. *P*-values indicate the probability that A' was significantly greater than .5. [a]Aggregate—reflecting the facially-triggered transference effect—was computed by reverse scoring aggressiveness, and computing the mean A' for the six trait judgments.

Table S2. Tendency to judge partner-similar (vs. yoked-similar) faces as possessing a particular trait, reflected by mean A′ scores for each of the six trait judgments and their aggregate (i.e., facially-triggered transference), for women and men, including only participants who did *not* express subjective awareness of the resemblance.

	Women's A'				Men's A'				Women's A' compared to men's	
	Mean	SE	*t*	*p*	Mean	SE	*t*	*p*	*t*	*p*
Trustworthy	0.67	0.05	3.62	0.0011	0.46	0.05	-0.94	0.3549	3.54	0.0016
Supportive	0.65	0.05	2.96	0.0051	0.56	0.05	1.37	0.1790	1.45	0.1636
Intelligent	0.61	0.05	2.27	0.0288	0.47	0.04	-0.61	0.5423	2.25	0.0331
Attractive	0.63	0.05	2.44	0.0193	0.56	0.05	1.26	0.2166	0.93	0.3602
Accepting	0.64	0.05	2.72	0.0095	0.51	0.05	0.20	0.8432	1.90	0.0711
Aggressive	0.45	0.04	-1.23	0.2241	0.55	0.04	1.26	0.2157	-1.76	0.0858
Aggregate[a]	**0.62**	**0.04**	**3.23**	**0.0025**	**0.50**	**0.04**	**0.07**	**0.9467**	**2.46**	**0.0218**

$N = 43$.

Notes. A' is a sensitivity measure adjusted for response bias. An A' of .5 reflects chance responding. An A' significantly greater than .5 reflects the tendency to judge partner-similar (vs. yoked-similar) faces as possessing the trait. *P*-values indicate the probability that A' was significantly greater than .5. [a]Aggregate—reflecting the transference effect—was computed by reverse scoring aggressiveness, and computing the mean A' for the six trait judgments.

Table S3. Tendency to judge partner-similar (vs. yoked-similar) faces as possessing a particular trait, reflected by mean A′ scores for each of the six trait judgments and their aggregate (i.e., facially-triggered transference), for women and men, statistically controlling for objective awareness of the resemblance.

	Women's A' compared to chance				Men's A' compared to chance				Women's A' compared to men's	
	Mean	SE	*t*	*p*	Mean	SE	*t*	*p*	*t*	*p*
Trustworthy[b]	0.65	0.04	3.96	0.0007	0.50	0.04	0.10	0.9243	1.95	0.0586
Supportive	0.63	0.04	3.00	0.0047	0.53	0.04	0.67	0.5073	0.79	0.4372
Intelligent[b]	0.63	0.04	3.50	0.0021	0.48	0.04	-0.49	0.6292	2.28	0.0287
Attractive	0.62	0.04	2.88	0.0063	0.59	0.04	1.99	0.0530	1.34	0.1872
Accepting[b]	0.57	0.05	1.36	0.2001	0.52	0.05	0.45	0.6606	2.04	0.0481
Aggressive	0.45	0.04	-1.26	0.2127	0.49	0.04	-0.32	0.7482	-0.93	0.3568
Aggregate[a,b]	**0.61**	**0.03**	**3.67**	**0.0014**	**0.52**	**0.04**	**0.46**	**0.6482**	**1.96**	**0.0572**

$N = 46$.

Notes. A' is a sensitivity measure adjusted for response bias. An A' of .5 reflects chance responding. An A' significantly greater than .5 reflects the tendency to judge partner-similar (vs. yoked-similar) faces as possessing the trait. *P*-values indicate the probability that A' was significantly greater than .5. [a]Aggregate—reflecting the transference effect—was computed by reverse scoring aggressiveness, and computing the mean A' for the six trait judgments. [b]For traits marked, the interaction between sex and $A'_{awareness}$ was statistically significant. So for these traits, separate linear mixed models were run for women and men to obtain estimated means and standard errors reported in the table. These means were then compared to chance by conducting one-sample *t*-tests.