Supporting Information

Lewis 10.1073/pnas.1308501110

Site Description

Despite the rapid rise of the internet as a romantic intermediary (1, 2), analyses of behavioral data from online dating sites still are rare. The site used for these analyses, OkCupid, is one of the more popular dating sites in the United States (3), although direct comparisons with other sites are difficult because membership data are not available. OkCupid does not market itself to any particular demographic group and therefore attracts a wide swath of the population. The absence of cost also eliminates a significant barrier to entry and likely increases the generalizability of these results. On the other hand, individuals who use a free online dating site might be considered less "serious" about finding a mate than those willing to pay for a subscription-based site, and, importantly, individuals who are particularly interested in finding a partner from one or another racial background likely will subscribe to one of the many available niche sites specifically tailored to this interest (4). Therefore, estimates of racial prejudice on OkCupid may be conservative relative to the general population.

Personal Profiles. Upon joining the site, users create a personal profile not unlike those on other popular social network sites, such as Facebook and LinkedIn. Rather than revealing their actual names or identifying information, however, users choose a screen name by which they will be identified on the site. The only information users are required to provide is their gender (female/male), sexual orientation (straight/gay/bisexual), relationship status (single/seeing someone/married), birth date, and five-digit ZIP code, all of which (except the full ZIP code) are visible to other users. Beyond this, users have the option of uploading any number of photographs to their profile (one of which serves as the primary photo used for identification around the site) and providing two types of personal information. First, users may describe themselves using several closed-ended prompts regarding their racial background, income, educational attainment, religious views, body type, and several other demographic, physical, and lifestyle descriptors. Second, users have the option of describing themselves in response to several open-ended essay prompts that serve as the main text of user profiles.

Contacting Others. Upon logging into the site (which is possible from a computer or smart phone application), users are greeted by a personalized welcome screen. Four general courses of action are available: (*i*) viewing and editing one's own profile; (*ii*) searching for, viewing the profiles of, and/or potentially contacting other site users through a variety of means; (*iii*) viewing one's inbox of messages received from other site users and potentially replying to these messages; and (*iv*) answering questions about one's romantic preferences (see below).

There are six methods through which users can locate other users on the site. Each method, in order of how commonly it is used (least to most common), is described below (5):

i) "Quickmatch" (1% of all profile views). OkCupid's most basic matching function simply recommends a single match upon clicking a button on the welcome screen. An abbreviated version of this person's profile (including some text, photos, and a match percentage) then is presented to the user, who has the option of indicating her level of interest and deciding whether to view that person's full profile. This process may be repeated as frequently as the user wishes.

ii) "Quiver" (5% of all profile views). OkCupid's Quiver feature is similar in function to Quickmatch, except that the site provides new recommendations automatically rather than in response to user-initiated queries. Users can view the recommendations in their Quiver at any time and indicate their level of interest in each recommendation; if a certain recommendation is dismissed, it is repopulated with a new recommendation after a period of some days.

iii) Seeing who visited one's profile (5% of all profile views). On their home page, users can view a list (including thumbnail photos and match percentages) of all users who recently visited their profile. Users may disable this feature, but they then no longer appear on this list on the home pages of others. (In other words, to view someone's profile without that person knowing, one must sacrifice the ability to know who has visited one's own profile.)

iv) Home page newsfeed (21% of all profile views). On their home page, users also are provided with a list that displays recent activity of other site members (somewhat akin to Facebook's "News Feed" function), such as when other members have uploaded a new photograph or edited something on their profiles. Thumbnail photographs of these users also are provided.

v) Viewing the profiles of others (23% of all profile views). Every time a user views the profile of another user, she is presented with a list (including thumbnail photos) of other users who are similar to the person currently being viewed, in whom the viewer may be interested.

vi) Search (45% of all profile views). The most common means of locating other users on OkCupid is to use the site's personalized search feature. Users have the option, first, of entering any number of search criteria that correspond to the closed-ended responses users provide on their profile and, second, of ordering search results any number of ways (e.g., by geographic proximity, match percentage, or how recently the user joined the site). Search results appear as a list of thumbnail photos, along with each user's screen name, match percentage, and a very brief excerpt from one of their open-ended essay responses.

Upon arriving at the profile of another user, regardless of which of the above methods is used, the person viewing the profile has the opportunity to send that user a message using the Web site's internal e-mail system.

Match Percentages. OkCupid differentiates itself from other Web sites not only through its absence of membership fees, but also through its unique approach to matching. After joining the site, users have the option of "improving their matches" by answering a virtually unlimited supply of questions about themselves and their mating preferences. The more questions a user answers, the more data the site acquires about the user and her preferences, and therefore the "smarter" the site becomes at recommending matches. The substance of these questions is extremely broad—ranging from the importance of sex vs. love, to one's views about homosexuality, to how likely one is to smile at a child walking down the street—and only rarely have to do with demographic preferences (presumably because one can easily realize these preferences using the search function). For each question, users must provide three responses, respectively indicating (*i*) their answer to the question, (*ii*) what responses to the question they will accept from a potential match, and (*iii*) how important the question is (options are "irrelevant," "a little important," "somewhat important," "very important," and "mandatory"). Users also have the option of skipping the current question and moving on to the next.

As a consequence of completing these questions, a "match percentage" appears on every profile a user visits. This percentage also appears virtually any time the site recommends a user via any of the methods described above. Match percentages are calculated based on the values of *i*, *ii*, and *iii* above for all questions that both site users have answered, using a straightforward (if nuanced) algorithm that OkCupid publicizes on its Web site (www.okcupid.com/help/match-percentages).

The Problem of Site Interference. If individual preferences and prejudices are to be inferred from patterns of interaction on an online dating site, then we must be sure these patterns reflect the autonomous choices of individuals rather than the external influence of the site (or rather, that site interference explains, at most, a negligible proportion of this behavior). There are three reasons to believe this is the case for OkCupid. First, match percentages, which constitute the backbone of OkCupid's approach to matchmaking, are calculated transparently based exclusively on user-provided preferences. In other words, the match percentage, like a sophisticated version of the search function, simply feeds back to the user what she herself has requested. Among the six possible ways one user can find another, only the Quickmatch and Quiver functions are relatively opaque in how these matches are generated; however, (*i*) these collectively account for only 6% of all profile views and (*ii*) there is every reason to expect that match percentages constitute the basis for these recommendations, as well.

Second, even when the site recommends a match via one of the above six methods, it still is the user's decision whether to view the full profile of that person and, even more importantly, whether to send that person a message. At least two prior teams of researchers studying online dating have taken the approach of examining messaging only conditional on profile views (6, 7). However, because race is a characteristic one can search for as well as a characteristic that is visible, this approach ignores perhaps the two most important stages of selection in which racial prejudice will operate. If the degree of out-group aversion among individuals from a certain racial background is strong enough, they will never search for or view the profiles of out-group members in the first place (8).

Third, even recognizing that the ultimate decision to contact someone is in the hands of the individual user—and recognizing that site recommendations essentially are a means of preselecting a pool of eligible partners on the basis of users' explicit preferences—it still is possible that the site interferes with these recommendations in some unknown way. For this interference to influence the findings of this paper, however, (*i*) such interference would have to be related directly or indirectly to racial background (interference that is unlikely to avoid detection); (*ii*) such interference would run directly contrary to the site's marketing strategy regarding its unique, transparent, user-driven approach to matching; and (*iii*) one would have to tell a complex story for how the site might be interfering in such a way as to spuriously produce the two central findings of this paper: that cross-race replies are more likely than cross-race initiations, and that receiving a cross-race initiation causes users to extend more cross-race initiations to other users in the future.

Study Population

The baseline population from which I selected the sample for this study consisted of all OkCupid users who self-identified as single, straight, living in the United States, and "looking for" short-term and/or long-term dating, and who joined the site between October 1 and November 30, 2010. Given that this paper focuses on racial prejudice in the United States; that it is reasonable to expect mating dynamics to vary substantially according to relationship status, sexual orientation, and interpersonal objectives on the Web site (other options included "new friends," "activity partners," "long-distance pen pals," and "casual sex"); and that enforcing a time window prevents the truncation issue described in the main text, these criteria seemed reasonable as a starting point for exploring racial dynamics in online dating. I also acquired the timestamp, sender, receiver, and length (in characters) of all messages sent among this study population from October 1, 2010, through December 15, 2010. I placed several additional restrictions on these data for the purposes of this paper.

Individual Data. Among the OkCupid users, 165,525 satisfied the individual criteria listed above. A long tradition of research has documented the importance of controlling for opportunity structures in any estimation of in- and out-group preferences; it is impossible to form relationships with individuals from another background if no such individuals are available in the first place (or if they live unreasonably far away to consider contacting). Although all users provide their five-digit ZIP code to OkCupid as a condition for site registration, I received access to only a truncated version of these codes to protect against the possibility of reidentification. Further, regional data were stripped entirely from individuals who lived in regions with comparatively few site members. Therefore, my first step was to eliminate from the baseline population the 3,143 users (1.9%) for whom regional data were not available.

Among the closed-ended demographic descriptors that appear on users' profiles, site users have the option to choose from several racial categories: Asian, Middle Eastern, black, Native American, Indian, Pacific Islander, Hispanic/Latin, white, or "other." It also is possible to select more than one category or not to select a category at all. For the sake of simplicity, as well as to ensure that the methodological tools of this paper (which require a minimum number of inter- and intragroup messages) could be used, I focused only on users who (*i*) provided data on their racial background, (*ii*) selected only one category of racial affiliation, and (*iii*) belonged to one of the five largest racial categories. Of the baseline population of 162,382 users now available, 140,395 (86.5%) selected at least one racial category. Of these, 130,159 (92.7%) selected only one racial category, and of these, 126,134 (96.9%) selected one of the five largest racial categories. This constituted the final study sample for these analyses, with the demographic composition described in Table S1.

Relational Data. I first limited the sample of messages to only those sent and received by one of the 126,134 users described above. Otherwise, some messages would be sent to or received from users outside the study sample, whereas a concrete network "boundary" is required for the network analytic techniques in this paper. Next, as described in the main text, I focused only on first messages and first replies. Further, I defined a "reply" as a response sent within 2 wk of the initial message. Two weeks seemed a conservative window within which to expect a reply if the romantic interest indeed was reciprocal. Because relational data were available through mid-December—but to prevent the same kind of truncation problem that I avoided at the start of the time window—I therefore considered initiation messages sent between October 1 and November 30 (the same interval in which new users were allowed to join) and responses sent through December 14. (Otherwise, if I had included initiation messages sent through mid-December, it would be impossible to distinguish a message that never received a reply from one that received a reply beyond the time window of available data.) Finally, because the focus of this paper is on romantic contact among heterosexual users, I eliminated all same-sex messages (which accounted for less than 1% of total messages among this sample).

Sample Representativeness. An important question is the extent to which online dating site users—and OkCupid users in particular—are representative of the broader population of singles in the United States. We know from previous research that age, education, and income are strongly associated with the likelihood of internet access (9), and online daters are younger, make less money, and are more likely to be used compared with nononline daters (10). However, once Sautter et al. (11) limited consideration to single internet users in the United States (i.e., people "at risk" of online dating), they found that no sociodemographic variables are significantly associated with the likelihood of internet dating. Comparing the specific sample of OkCupid users in these analyses with the broader population of unmarried, internet-using adults in the United States from the relevant five racial backgrounds, the sample contains fewer blacks and Hispanics, more whites, and slightly more men than the general population. Users in the sample also are more educated, more likely to be in their 20s and 30s, and make slightly less money than we would expect by chance (although only household income data are available for the broader population, which makes the latter difference unsurprising).[†]

Methodological Details

Exponential Random Graph Modeling. Exponential random graph (ERG) modeling is increasingly common in mainstream social scientific literature (12–15). However, the computational demands required to estimate these models place hard limits on a combination of model complexity and the size of the population over which such models may be estimated. Consequently, most applications of ERG models involve networks on the order of, at most, hundreds or a few thousand nodes (if not far fewer), and applying this method to the entire study sample used here is far outside the realm of possibility. Fortunately, it also is outside the realm of desirability; site users interested in dating rarely will initiate contact beyond a certain geographic radius (as noted in the main text), so it is hardly reasonable to consider the entire study sample as the practical opportunity structure for all users. Therefore, I estimated a separate model for each two-digit ZIP code region[‡] (for all regions with more than 1,000 users represented in the sample) and assessed the overall importance of each micromechanism by looking at summary statistics of parameter estimates across all models. This technique also was used by Goodreau et al. (12) in their study of homophily and triadic closure in American high schools. The full set of parameter estimates from all models is presented in Table S2. As noted in the main text, a primary strength of these models is their great degree of flexibility; ERG models can, in principle, incorporate any number of optional parameters that correspond to virtually any possible reason one site user might contact another. Here, however, I model online messaging behavior solely as the joint product of gender, racial background, and reciprocity to present results that are as parsimonious as possible and tailored to the primary questions of this study.

A few notes of caution about these models are in order. First, although I included only regions with more than 1,000 users represented in the sample (collectively accounting for 81.3% of the entire study sample) to increase the likelihood that each region would contain users from all five racial backgrounds, it is clear from Table S2 that the issue of "cell size" still was a problem. Specifically, even if there were more than one individual from a particular racial background in a given region (or, more precisely, at least one male and at least one female), it is not possible to estimate parameters for racial matching and the interaction of racial matching and reciprocity unless there is at least one same-race tie and at least one same-race reciprocal tie among those users, respectively. Consequently, it is important to recognize that the box plot for each effect in Fig. 2 pertains only to regions in which a parameter estimate for that effect could actually be obtained (ranging from 100% of all regions for racial matching among white users, e.g., to 18% of all regions for the interaction of racial matching and reciprocity among Indian users).

Second, despite the recent advances in ERG modeling used here (including simulation-based estimation), ERG models traditionally have been plagued by a problem known as "degeneracy" when a model is poorly specified and parameter estimates for that model produce empirically implausible networks, such as a graph with no ties at all or with all nodes connected to all others (16). To ensure that degeneracy was not a problem here, I used the recommended technique in the literature for assessing the fit of ERG models: running a large number of network simulations (here, 100) based on the final parameter estimates for each model, and comparing global properties of these simulated networks to the global properties of the actually observed social network (17). A typical goodness-of-fit output appears in Fig. S1, reflecting model fit for the out-degree distribution (the quantity of messages sent by each user), in-degree distribution (the quantity of messages received by each user), and geodesic distance distribution (the quantity of network steps between any two users). Although, in general, global properties of simulated networks departed substantially from global properties of the observed networks, this is not surprising given the simplicity of models used here (e.g., we would expect that a great deal of variation in in-degree, out-degree, and geodesic distance would be explained by other factors, such as individual attractiveness). Further, the fact that the fits of these models are as good as they are is in some sense quite remarkable and reflects how strongly site behavior is, in fact, patterned simply by the joint dynamics of gender, reciprocity, and racial background. Most importantly, however, virtually none of the models estimated in this paper shows any sign of degeneracy; in other words, almost all models successfully "converged" (the sole exception is noted in Table S2).

Third, it is important to emphasize that all effects in each model are estimated controlling for all other effects in the model, and should be interpreted in this light. In other words, it is not necessarily the case that dating site users prefer to reciprocate ties to individuals from a different racial background in an absolute sense, but rather that the log odds of reciprocating a tie to someone from a different racial background (as captured by the matching*reciprocity interaction effects) generally are greater than (for Indian, Asian, and Hispanic users) or equal to (for white and black users) what we would expect by chance conditional on the baseline tendency toward reciprocity and the great degree of in-group bias displayed by most users when initiating contact. In other words, matching*reciprocity interaction coefficients should not be interpreted out of the context of the distinct matching and reciprocity effects that also are estimated in each model.

To illustrate this point further, Fig. S2 presents the same results as in Fig. 2, but in an entirely different fashion. Whereas Fig. 2 presents summaries of the actual model coefficients across all regions, Fig. S2 presents summaries of the log odds of all possible messaging scenarios across all regions. In other words, whereas Fig. 2 disentangles the magnitude of the various effects included in each ERG model and presents these effects side by side, Fig. S2 quantifies the direct implications of these effects for the log odds of all possible messaging scenarios among site users. All results in Fig. S2 therefore are derived directly from the data featured in Fig. 2 (and presented in full in Table S2) based on very straightforward calculations. To determine the log odds of any given

[†]Comparisons are based on the October 2010 Internet Use Supplement of the Current Population Survey.

[‡]A helpful map representing the geographic size and location of all 2-digit ZIP Code regions can be found at http://benfry.com/zipdecode/.

messaging scenario, one simply adds all relevant parameter estimates to describe that scenario. So, for instance, the log odds of a female site user from one racial background initiating contact with a male site user from a different racial background ("female initiation–cross-race") are simply equivalent to the "density" coefficient, because no other model terms describe this scenario. The log odds of a male black user initiating contact with a female black user ("male initiation–both black") are equivalent to the density coefficient added to the "female-receiver" coefficient and the "black-matching" coefficient. Also, the log odds of a male Indian user replying to an Indian female user ("male reply-both Indian") are equivalent to the density coefficient added to the "reciprocity" coefficient, the female-receiver coefficient, the "Indian-matching" coefficient, and the "Indian-matching*reciprocity" coefficient (to choose one of the most complex possible scenarios represented by this model).

Finally, the advantage of ERG models is their ability to statistically disentangle and quantify the importance of various underlying mechanisms in producing overall network structures (13), and particularly to compare patterns of initiation with patterns of reply. Their primary disadvantage in the context of this study—and the reason affiliation indices are still helpful—is that because these models were designed for cross-sectional data, the dimension of time must be suppressed in model results.[§] In other words, the network I model is merely a "snapshot" of all interactions over the course of the study period, with the restrictions described in the main text. However, the consequences of this fact for the findings presented in the main text actually are quite minimal. The reasons for this are twofold. First, although ERG models are indeed designed for cross-sectional data, they assume that these data were produced by dynamic underlying processes; in fact, the simulation-based techniques used to estimate these models respect this underlying dynamism. Second, the upshot of this consideration is that relatively little information relevant to this study actually is lost; in fact, the only practical ramification of the suppression of time for the current paper is that the ERG models, in circumstances in which both an initiation and a reply are present between two users, cannot determine which of the two users initiated the exchange. For this reason, three-way interactions among race, gender, and reciprocity are not possible. However, as is evident from Fig. S2, this does not mean that gender is altogether absent from patterns of reciprocity, because these patterns still consider the female-receiver effect (which incorporates females' tendency to receive initiations as well as replies).

Coarsened Exact Matching. To help understand the implementation of my counterfactual approach, a visualization of the partitioning of the study period into control, treatment, and outcome periods is presented in Fig. S3 (20). Although I present findings in the main text in terms of the average quantity of interracial exchanges initiated to facilitate interpretation of results, I replicate all findings from the matching analyses in Table S3 in their original, logged form (including SEs, significance levels, and the quantity of observations in each analysis). It is important to note that all effects presented in the main text are conservative for two reasons. First, I estimated all models without controlling for the additional imbalance that remained in the matched data as a consequence of the fact that the treatment and control groups were matched coarsely (instead of exactly) on some variables. The reason for this omission is, first, to facilitate interpretation of the causal effect (as simply the difference in mean values of the outcome variable between treatment and control groups, instead of the average difference between groups controlling for other factors) and second, in one case (the model for Asian women) the model with added controls failed to converge. However, in nearly all instances in which I replicated these models but also included controls, the average treatment effect on the treated increased in magnitude compared with models without controls (see *Supplementary Analyses* below).

The second reason these findings are conservative is that the outcome variable—like all other relational variables in these analyses (including prior exchanges initiated, prior interracial exchanges initiated, and the treatment variable itself)—was measured only with respect to ties sent and received among the original study sample of 126,134 users. In other words, a strict interpretation of the baseline effect (Fig. 3*A*) is not that users who received the treatment initiated an average of 0.038 new interracial exchanges per person than they would have otherwise (0.141 minus 0.103), but rather that users who received the treatment initiated an average of 0.038 new interracial exchanges with other users in the study sample. Users in the treatment group, like all other users in the sample, were almost certainly interacting with other individuals on the site besides simply those users who had also joined the site recently, expressed that they were interested in "dating," and so forth. Additionally, this causal estimate does not include interracial contact with any site user who belonged to one of the other racial categories on the site that were not included in these analyses. Naturally, according to this logic, it also is possible that individuals who were considered part of the control group for these analyses in fact received an interracial message during the treatment period from a user who was not herself part of the study sample; therefore, such individuals were not counted as receiving the treatment. However, this again would create a conservative bias in results. It is possible that some users in the control condition "actually" received the treatment (which would lessen the gap in outcomes between treatment and control groups, not widen it), but by definition no one in the treatment condition could have actually been untreated.

A final detail of these models is worth noting. For almost all analyses, I maximized available data by allowing multiple cases to serve as controls for each treatment. The sole exception to this approach is the analysis presented in Fig. 4*C* ("scope of effect"). Here, I estimated the average treatment effect on the treated for two distinct outcome variables: the quantity of interracial exchanges initiated with users from the same racial background as the treatment sender, and the quantity of interracial exchanges initiated with users from a different racial background as the treatment sender. However, both these outcomes are defined in relation to the racial background of the treatment sender, although by definition the sender of the treatment does not exist for the control group and, therefore, it is unclear how the counterfactual outcome should be measured. To address this problem, rather than allowing multiple cases to serve as controls for each treatment, I required that each treatment case be assigned to precisely one control case and I personalized measurement of the counterfactual outcomes for that control case in precisely the same way as I measured the actual outcome for the treatment case. Specifically, for instance, if a certain treatment case was an Asian male who received an interracial message (treatment) from a black female, I defined this person's outcome variables as the quantity of future exchanges initiated with other black females (specific effect) and the quantity of future exchanges initiated with Indian, Hispanic, or white females (generalized effect), and I measured the outcome variables for this person's control case (another Asian male who did not receive a treatment) identically (i.e., the quantity of future exchanges initiated with black females, for the specific

[§]Recently, Krivitsky and Handcock (18) developed an extension of ERG modeling for dynamic data. However, these models require that data be recorded in discrete snapshots (rather than continuous time) and that we can properly speak of relationship "formation" and "dissolution," whereas in the case of online dating, an instance of two people who suddenly stop messaging could equally plausibly reflect a relationship "failure" (i.e., one or the other lost interest) as a "success" (i.e., the interaction transitioned offline). One promising direction for future work is the relational event framework developed by Butts (19).

effect, and with Indian, Hispanic, or white females, for the generalized effect). I also limited this analysis to users in the treatment group who received one and only one interracial message during the treatment period to avoid ambiguous situations in which site users were "treated" by exposure to individuals from more than one racial category. Such users accounted for the majority (78%) of treated cases.

Supplementary Analyses

I conducted five sets of supplementary analyses to clarify and confirm results presented in the main text. I present the nature and results of these analyses below.

"Polite Rejections." An important alternative explanation for why cross-race messages are more common (or, equivalently, same-race messages are less common) for replies than for initiations is that dating site users fear being perceived as racist; therefore, they are more likely to send a polite rejection (e.g., "thanks, but no thanks") to someone from a different racial background than to someone from the same racial background rather than simply ignoring the initial message altogether.[¶] To test for this possibility, I first compared the distributions of message length (in characters) between cross-race replies and same-race replies, because it seems reasonable to expect that a polite rejection would tend to be shorter in length than a reply conveying genuine interest. As presented in Fig. S4, it is indeed the case that relatively short messages are slightly more common among cross-race replies than among same-race replies. Second, to assess whether this difference is driving results, I replicated the analyses presented in Fig. 1 four times—once for each reply length quartile—and focused primarily on the difference between rates of same-race initiation and same-race response (i.e., the figures presented on the diagonal of Fig. 1 *C* and *F*). As is clear from Fig. S5, the same-race affiliation index is always higher for initiations than for replies (i.e., the "difference" bar is negative) when we focus only on messages in the shortest quartile (1–37 characters), regardless of gender and racial background. However, this difference almost always exists for replies in each of the three other message length quartiles as well (and in fact, in the case of Hispanic men, black women, and Hispanic women, the difference between rates of same-race initiation and same-race reply widens the longer the message). Further, the only instances in which same-race affiliation indices were higher for replies than for initiations were for Indian men (when the length of their reply was in the second, third, or fourth quartile) and Asian women (when the length of their reply was in the second quartile), the first of which can already be observed in Fig. 1*C* in the main text. In other words, although it is clear that the precise difference between the likelihood of a same-race initiation and a same-race reply indeed depends on the length of the reply—and in some cases, this difference is larger the shorter the response (particularly for white dating site users, who might be especially concerned about perceptions of racism)—the difference between initiations and replies occurs almost universally regardless of reply length, and there is little reason to fear that polite rejections are driving this finding.[‖]

Same-Race "Treatment." Turning next to the matching analyses, one alternative explanation is that the apparent effect of the treatment does not stem from having been contacted by someone from a different racial background but from having been contacted by anyone at all. In other words, one can imagine a scenario in which receiving an interracial message per se does not make dating site users more likely to reach across racial boundaries in the future, but that receiving a message from anyone makes dating site users more likely to initiate contact with anyone (including individuals from a different racial background) in the future (because perhaps receiving a message from anyone boosts site users' confidence, or perhaps users who receive messages are more likely to remain on the site longer and therefore more likely to send out messages in the future). To check for this possibility, I conducted two additional analyses.

First, if the effect documented in this paper is an artifact of receiving a tie from anyone, then we would expect that users who receive a same-race version of the treatment also would initiate more interracial exchanges in the future. We therefore can consider three categories of site users: (*i*) users who received at least one cross-race message during the treatment period (i.e., the actual treatment in this study), (*ii*) users who received only same-race messages during the treatment period, and (*iii*) users who did not receive any messages at all during the treatment period. I find that users in category *ii* indeed initiate significantly more interracial exchanges during the outcome period than users in category *iii* ($P < 0.01$). However, the size of this effect is smaller than the size of the effect when I compare users in category *i* with users in category *iii*, which also is statistically significant ($P < 0.01$; Fig. S6*A*).

Second, to directly assess the difference between the two treatment groups, I dropped all individuals in category *iii* and instead estimated the average treatment effect on the treated when I matched all users in category *i* with at least one user in category *iii*. I found that users who received an interracial treatment initiated significantly more interracial exchanges during the outcome period than did users who received a same-race treatment ($P < 0.001$; Fig. S6*B*). In other words, simply receiving a message from anyone indeed causes dating site users to initiate more cross-race exchanges in the future, likely because receiving a message from anyone makes users more likely to initiate exchanges with anyone in the future. However, the results I present in this paper cannot be reduced to such an effect; the fact that the sender of the treatment message is from a different racial background than the receiver is indeed causally consequential.

Controlling for Remaining Imbalance. A second alternative explanation for these findings is that although I matched treatment and control groups exactly according to gender, racial background, and two-digit ZIP code region, I matched these two groups only "coarsely" according to previous quantity of initiation messages sent, previous quantity of cross-race initiation messages sent, and account age. Therefore, the apparent difference between treatment and control groups could be explained simply by the remaining imbalance between the two groups with respect to the latter three variables. To check for this possibility, I ran the same model presented in Fig. 3*A*, but I also included the original versions of these three variables as controls. Individuals in the treatment category continued to initiate significantly more interracial exchanges during the outcome period than individuals in the control category ($P < 0.001$); in fact, the effect size increased, rather than decreased, in magnitude (Fig. S7). I repeated this robustness check for all other analyses presented in the main text (results not pictured). Inclusion of these controls almost always resulted in an increase in the magnitude (and a decrease in the P value) of the treatment effect; in no case did an effect that was statistically significant previously cease to be so, and in two cases an effect that was not statistically significant previously became statistically significant with the inclusion of controls (the case of female site users whose account was 9–16

[¶]I thank an anonymous reviewer for this suggestion.

[‖]Unfortunately, it is not possible to replicate the ERG analyses presented in Fig. 2 using the different subsets of replies examined in Fig. S5, because there would be too few replies in the data for most two-digit ZIP code regions. (For instance, even considering replies of all lengths, we can see from Table S2 that it was not always possible to obtain parameter estimates for the interaction of racial matching and reciprocity because the relevant type of tie—a reply sent between two site users from the same racial background—was not present in the data.)

d old, $P < 0.001$, and the case of site users who had received an interracial message previously, $P < 0.001$).

The Problem of Unobservables. The final and potentially most problematic alternative explanation for these findings is that the sample is not balanced on unobserved characteristics associated with both the treatment and the outcome. In other words, there may be one or more individual characteristics that make certain dating site users both (*i*) more likely to attract interracial messages and (*ii*) more likely to send interracial messages, such that the effect of the treatment is not truly causal. This is a different problem from the alternative explanation above, which had to do with lack of balance on observed characteristics and in which the degree of imbalance was bound strictly by the level of coarsening; it also is the primary advantage of experimental research over observational research (i.e., randomized assignment to the treatment condition ensures balance on both observed and unobserved characteristics).

There are two reasons the findings of this paper are unlikely to be a spurious consequence of unobserved differences. First, in order for unobserved differences to explain the effect of the treatment on the treated, not only would these differences have to be correlated with both the treatment and the outcome variables, but the differences must not be highly correlated with any of the observed variables on which treatment and control groups were matched. Given that one of the observed control/matching variables is prior degree of interracial contact (i.e., quantity of interracial exchanges initiated before the treatment period, or a pretreatment version of the outcome variable), such an explanation is very unlikely. In other words, the unobserved characteristic would somehow have to be associated with the likelihood of receiving an interracial tie during the first week of November and the likelihood of sending interracial ties during the second week of November, but not associated with the likelihood of sending interracial ties during October. It is difficult to think of any such characteristic.

Second, if there is in fact some unobserved difference between treatment and control groups, then we would expect that differences in outcomes between the two groups should persist over time. In other words, if the effect of the treatment on the treated is not genuinely causal, then there is no reason we should expect the strength of this "effect" to diminish over time. To assess this possibility, I reanalyzed these data using two alternative versions of the outcome variable: one measured during the third week of November (two weeks after the treatment) and one measured during the fourth week of November (three weeks after the treatment). The results of this analysis are presented in Fig. S8, in which I also control for the remaining imbalance between treatment and control groups on observed dimensions, as in Fig. S7 (a decision that influenced the size of the treatment effect, but not its statistical significance). I find that, contrary to what we would expect if unobserved differences were motivating the treatment effect, the effect of the treatment on the treated decreases to statistical insignificance as early as 2 wk beyond the treatment. In other words, there is convincing evidence that the causal effect documented in this paper is indeed veracious, and that it endures for only a week beyond the time of treatment. After this point, treatment and control groups again become statistically indistinguishable with respect to their quantity of interracial initiation.

An additional and suggestive observation from the above exercise is worth noting. This observation is that although the effect of the treatment decreases to statistical insignificance as early as the third week of November—and remains statistically insignificant during the fourth week of November—the size of the (insignificant) effect actually increases from week 3 to week 4. The implications of this difference are twofold. First, given that Thanksgiving falls on the fourth Thursday of November, it is possible that this slight increase in effect size relates to unusual activity surrounding the holiday. For instance, site users may be particularly interested in finding a mate and, therefore, unusually susceptible to the effect of their prior treatment. Second, however, because the size of the treatment effect decreases from the second week of November to the third week of November—and only then increases from the third week of November to the fourth week of November—there is little reason to fear that any such possible "contaminating" effects of Thanksgiving may have influenced the focal outcome of this study, which was measured during the second week of November.

Understanding the Causal Mechanism. In addition to addressing alternative explanations for my findings, I conducted two final analyses to better understand the causal mechanism driving results and the specific subpopulations to whom this mechanism is applicable. First, we know from the main text that site users who receive an interracial message are more likely to initiate new interracial exchanges in the future. However, this effect might be driven by one or more of the following three mechanisms: First, as a consequence of the treatment, site users might be more likely to consider (i.e., view the profiles of) individuals from a different racial background; second, site users might be more likely to contact a prospective partner from a different racial background, conditional on viewing that person's profile; or third, site users might be more likely to contact prospective partners from a different racial background without viewing these individuals' profiles at all. In Fig. S9, therefore, I present results from a replication of the central analysis in the main text (presented in Fig. 3*A*), except that I redefine the outcome variable in three different ways corresponding to the three different scenarios above: first, as the quantity of cross-race profile views (Fig. S9*A*); second, as the cross-race initiation rate conditional on a cross-race profile view (Fig. S9*B*); and third, as the quantity of cross-race initiations that were not preceded by a profile view (Fig. S9*C*). As in the original analysis, I also matched treatment and control cases coarsely based on pretreatment overall and race-specific versions of the outcome variable (e.g., the total quantity of profile views and the total quantity of cross-race profile views, the overall initiation rate and the cross-race initiation rate); I measured all outcomes during the outcome period; and, although I continued to use negative binomial regression for the first and third analyses, I used a binomial generalized linear model with a logit link function for the second analysis (because the outcome variable, contact rate, is measured as a proportion rather than a count). Results clearly demonstrate that the finding in the main text is driven by the first and third mechanisms above. In other words, it is not the case that site users who receive the treatment are more likely to initiate contact with prospective partners from a different racial background after viewing these users' profiles; in fact, after the treatment they are significantly less likely to do so ($P < 0.05$). However, this lower rate of contact conditional on a profile view is overcompensated by the fact that site users who receive the treatment are more likely to view the profiles of prospective partners from a different racial background in the first place ($P < 0.001$). They also are more likely to contact individuals from a different racial background without viewing these individuals' profiles at all ($P < 0.001$). In other words, the causal effect of receiving a cross-race message is that it leads dating site users to "consider" (and also to "blindly" contact) prospective partners they would not have considered (or blindly contacted) otherwise, lending additional support to the "preemptive discrimination" interpretation proposed in the main text.

Second, we know (from Fig. 3) that the treatment effect on the treated varies in magnitude and statistical significance across men and women from different racial backgrounds who have been members of the Web site for shorter vs. longer periods. However, we also know (from Fig. 4*B*) that the treatment produces a causal effect only when the recipient replies to the treatment message. Therefore, how do the subgroup results from Fig. 3 change when we replicate analyses separately for site users who did and did not reply to the treatment message, as in Fig. 4*B*? Results from these replications are presented in Fig. S10 (in which, unlike Figs. 3 and 4*B* but like many other analyses in *SI Materials and Methods*, results are presented in terms of the actual parameter estimate and confidence interval for the treatment effect, not the difference in the quantity of interracial initiations between treatment and control groups). Once we distinguish between those who did and did not reply to the initial treatment message, it is clear that in many—but not all—circumstances, the observed causal effect in Fig. 3 is indeed driven by the subgroup of users who replied to the treatment message; in fact, if we focus only on the users who replied to the treatment message, the effect of the treatment is statistically significant for many more subcategories of users. All women who replied to the treatment message, regardless of account age, initiated significantly more interracial exchanges during the outcome period than women who did not receive a treatment message ($P < 0.01$). Almost all women who replied to the treatment message, regardless of racial background, initiated significantly more interracial exchanges during the outcome period ($P < 0.05$) (the only exception is Indian women—although the confidence interval for these users also is unusually large because of their small representation in the sample). Among men who replied to the treatment message, there is a positive, significant effect of receiving the treatment for men who joined the Web site in the past 1–8 d ($P < 0.001$), Asian men ($P < 0.01$), and Hispanic men ($P < 0.01$).

What explains this finding, and what explains the heterogeneity in this effect across individuals from different backgrounds? Two explanations are possible. One explanation is that there is a subset of site users for whom (for unknown reasons) racial boundaries are present but unusually fragile; in some sense, all it takes is a "push" from an external source to diminish the bias that previously prevented these individuals from considering cross-race prospects. This external push (i.e., receiving cross-race romantic interest) then causes these individuals both to reply to the original treatment message and to consider more interracial prospects in the short-term future.

The second explanation (note that the two are not mutually exclusive) is that the act of replying to the original treatment message itself carries causal weight. In other words, some subset of individuals who receive a treatment message reply to this message. It might be that something is unique about those circumstances when the reply occurs—whether an unobserved characteristic of the sender, an unobserved characteristic of the receiver, an unobserved characteristic of the message, or a particularly high level of compatibility between the two users—or it might be that these replies are more or less distributed by chance. Either way, it is then this very act of replying that temporarily reduces the level of preemptive discrimination these users display in the future—whether because the reply instigated a longer chain of interaction, and perhaps even a date, resulting in greater openness to considering future romantic prospects from that same racial background (21), or because users, observing their own reply, conclude that they must be open to considering individuals from that background and alter their subsequent search behavior accordingly (22).

If this is the case, however, why does this effect apply only to certain subsets of site users? All women, regardless of account age and racial background (with the exception of Indian women noted above), display the effect; yet, men display the effect only if they have joined the site relatively recently. If indeed it is the case, as suggested in the main text, that the treatment effect is driven by a change in site users' perceptions of others' racial prejudice—and therefore a change in site users' perceptions of their own likelihood of "success" with individuals from a different racial background—then we might expect exactly this difference given baseline gender dynamics in online dating. Women know their likelihood of receiving a reply from anyone is very high, whereas men know the vast majority of their advances will be ignored. In other words, even if men are persuaded that their chances with cross-race dating prospects are better than they originally thought, as long as they still think their chances with same-race prospects are better, they may continue to statistically discriminate in favor of the latter—unless they have joined the site relatively recently and therefore (*i*) have not yet learned how rare it is for men to receive a reply and (*ii*) are particularly impressionable regarding site norms and perceive that interracial contact is more common than it actually is. Meanwhile, the effect also is absent among black men, for whom the odds of receiving a reply from anyone are especially low, regardless of the racial background of the recipient (Fig. 1); Indian men, for whom the confidence interval again is especially large; and white men, for whom the effect of the treatment actually is significant only for those who do not reply to the treatment message. However, this likely is just a statistical artifact, given that the magnitude of the effect for those who do and do not reply to the treatment is quite similar, but the confidence interval is slightly larger among the former.

1. Rosenfeld MJ, Thomas RJ (2012) Searching for a mate: The rise of the internet as a social intermediary. *Am Sociol Rev* 77(4):523–547.
2. Cacioppo JT, Cacioppo S, Gonzaga GC, Ogburn EL, VanderWeele TJ (2013) Marital satisfaction and break-ups differ across on-line and off-line meeting venues. *Proc Natl Acad Sci USA* 110(25):10135–10140.
3. Paumgarten N (2011) Looking for someone: Sex, love, and loneliness on the Internet. *New Yorker*. Available at www.newyorker.com/reporting/2011/07/04/110704fa_fact_paumgarten. Accessed May 4, 2013.
4. Finkel EJ, Eastwick PW, Karney BR, Reis HT, Sprecher S (2012) Online dating: A critical analysis from the perspective of psychological science. *Psychol Sci Public Interest* 13(1): 3–66.
5. Piskorski MJ *Social Strategy: Why Social Media Platforms Work and How to Leverage Them for Competitive Advantage* (Princeton Univ Press, Princeton), in press.
6. Hitsch GJ, Hortaçsu A, Ariely D (2010) What makes you click? Mate preferences in online dating. *Quant Mark Econ* 8:393–427.
7. Skopek J, Schulz F, Blossfeld H-P (2011) Who contacts whom? Educational homophily in online mate selection. *Eur Sociol Rev* 27(2):180–195.
8. Lin K-H, Lundquist J Mate selection in cyberspace: The intersection of race, gender, and education. *Am J Sociol*, in press.
9. DiMaggio P, Hargittai E, Neuman WR, Robinson JP (2001) Social implications of the Internet. *Annu Rev Sociol* 27:307–336.
10. Madden M, Lenhart A (2006) *Online Dating: Americans Who Are Seeking Romance Use the Internet to Help Them in Their Search, but There Is Still Widespread Public Concern About the Safety of Online Dating* (Pew Internet and American Life Project, Washington).
11. Sautter JM, Tippett RM, Morgan SP (2010) The social demography of internet dating in the United States. *Soc Sci Q* 91(2):554–575.
12. Goodreau SM, Kitts JA, Morris M (2009) Birds of a feather, or friend of a friend? Using exponential random graph models to investigate adolescent social networks. *Demography* 46(1):103–125.
13. Wimmer A, Lewis K (2010) Beyond and below racial homophily: ERG models of a friendship network documented on Facebook. *AJS* 116(2):583–642.
14. Srivastava SB, Banaji MR (2011) Culture, cognition, and collaborative networks in organizations. *Am Sociol Rev* 76(2):207–233.
15. Papachristos AV, Hureau DM, Braga AA (2013) The corner and the crew: The influence of geography and social networks on gang violence. *Am Sociol Rev* 78(3): 417–447.
16. Handcock MS (2003) Statistical models for social networks: Degeneracy and inference. *Dynamic Social Network Modeling and Analysis*, eds Breiger R, Carley K, Pattison P (National Academies Press, Washington), pp 229–240.
17. Hunter DR, Goodreau SM, Handcock MS (2008) Goodness of fit of social network models. *J Am Stat Assoc* 103(481):248–258.
18. Krivitsky PN, Handcock MS A separable model for dynamic networks. *J R Stat Soc B*, in press.
19. Butts CT (2008) A relational event framework for social action. *Sociol Methodol* 38: 155–200.
20. Harding DJ (2003) Counterfacutal models of neighborhood effects: The effect of neighborhood poverty on dropping out and teenage pregnancy. *Am J Sociol* 109(3):676–719.
21. Allport GW (1954) *The Nature of Prejudice* (Addison–Wesley, Reading, MA).
22. Festinger L, Carlsmith JM (1959) Cognitive consequences of forced compliance. *J Abnorm Psychol* 58(2):203–210.

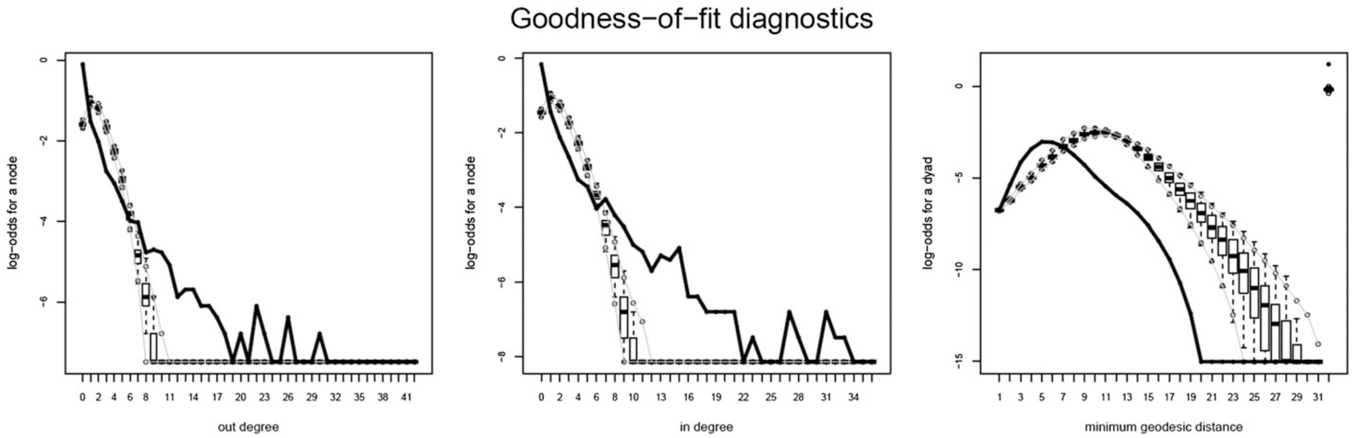

Fig. S1. Goodness-of-fit plots for the ERG model of ZIP code region 75xxx (n = 1,780). Solid lines represent observed network distributions, and box plots represent summaries of the same distributions from 100 simulated networks.

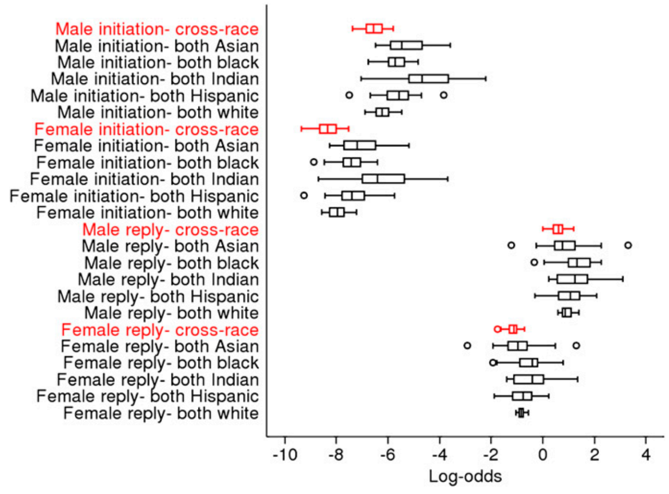

Fig. S2. Direct comparisons of the log odds of all possible messaging scenarios, derived from ERG model results summarized in Fig. 2 of the main text (and presented in full in Table S2). Cross-race scenarios are presented in red to facilitate comparisons with the same-race scenarios beneath them.

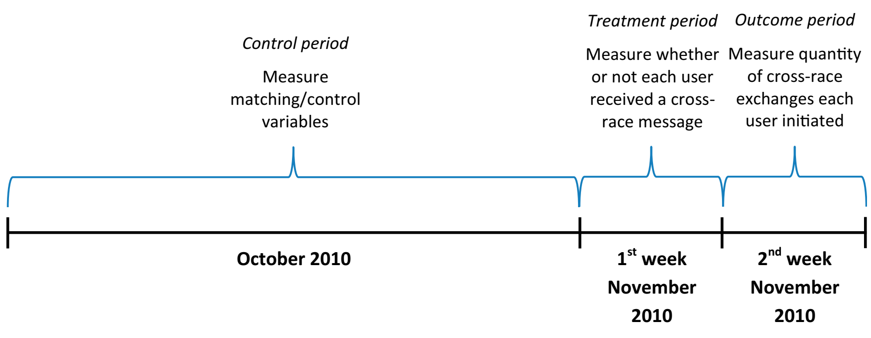

Fig. S3. Identifying the causal effect of receiving a cross-race message using a counterfactual framework. Matched treatment and control groups are identical (or coarsely identical) on observed characteristics before the treatment period. The difference in outcomes between matched treatment and control groups is the effect of the treatment on the treated.

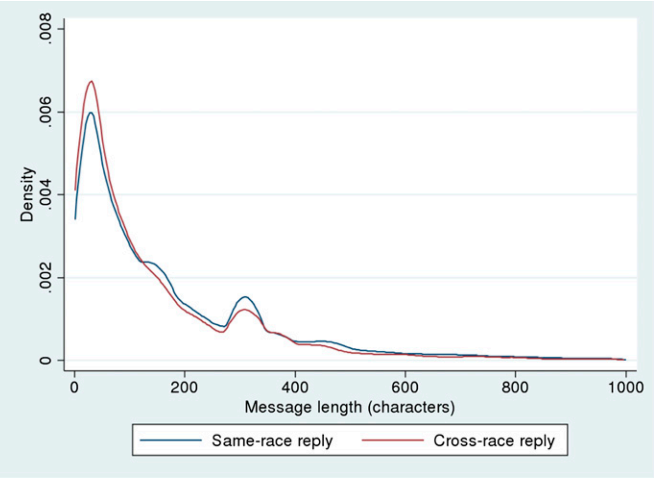

Fig. S4. Kernel density plots comparing the message length distribution for same-race replies (blue) and cross-race replies (red). To enhance visibility, only pictured are replies less than 1,000 characters in length, which account for 98.4% of all replies.

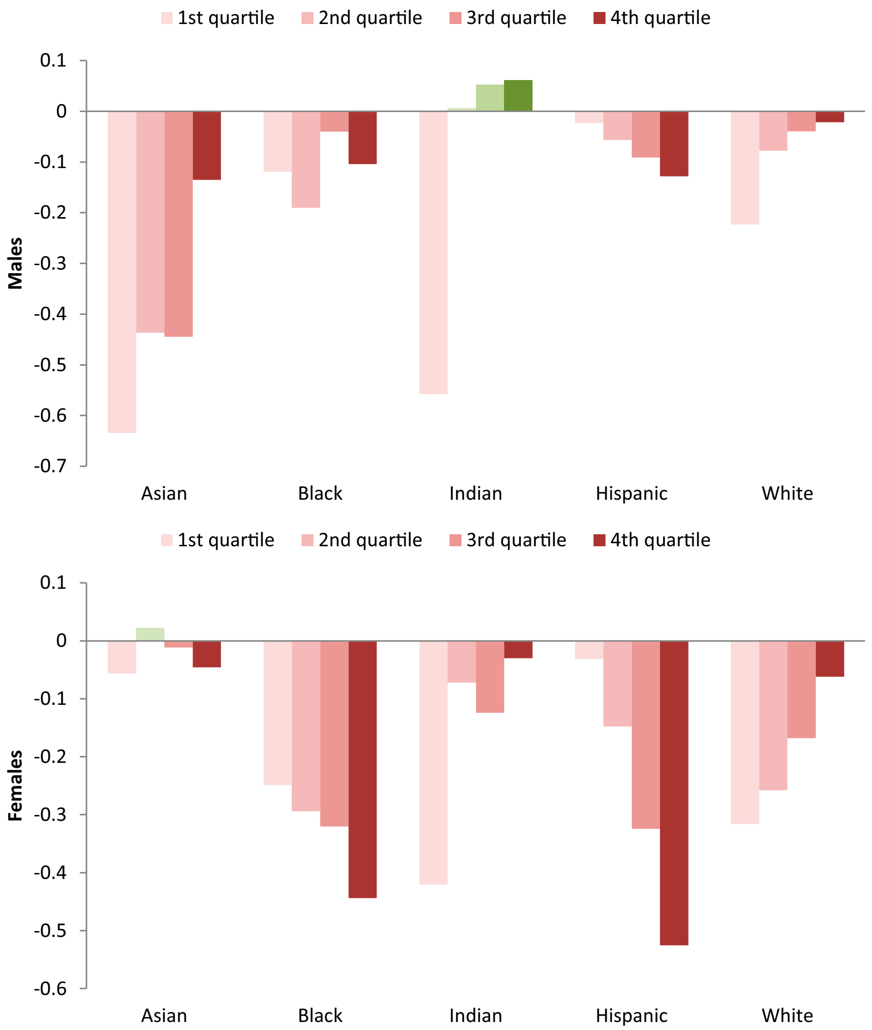

Fig. S5. Replication of Fig. 1 *C* and *F* from the main text across replies of different lengths, focusing only on same-race indices. In other words, this figure represents only the quantities featured on the diagonal of Fig. 1 *C* and *F*—the difference in same-race affiliation indices between patterns of initiations and patterns of replies—considered separately for replies in the first quartile (1–37 characters), second quartile (37–104 characters), third quartile (104–252 characters), and fourth quartile (252–10,182 characters) of message length.

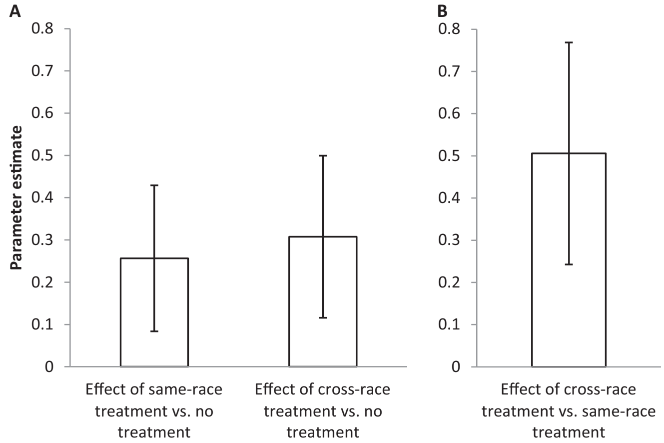

Fig. S6. Robustness check to ensure that the treatment effect is not simply an artifact of receiving a message from anyone. First, I compared the effect of receiving only same-race messages during the treatment period (relative to those who did not receive any messages) with the effect of receiving at least one cross-race message during the treatment period (relative to those who did not receive any messages) (*A*). Next, I assessed the effect of receiving at least one cross-race message during the treatment period relative to those who received only same-race messages (*B*).

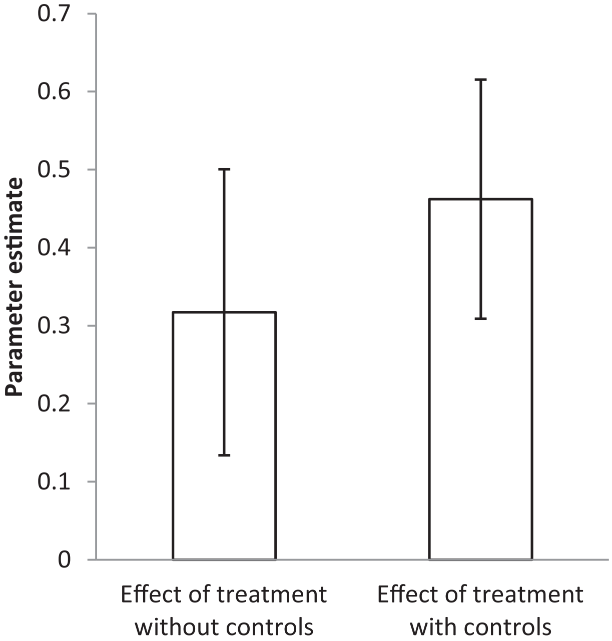

Fig. S7. Average treatment effect on the treated, estimated without and with additional controls (quantity of prior exchanges initiated, quantity of prior interracial exchanges initiated, account age) to account for remaining imbalance between treatment and control groups on observed dimensions.

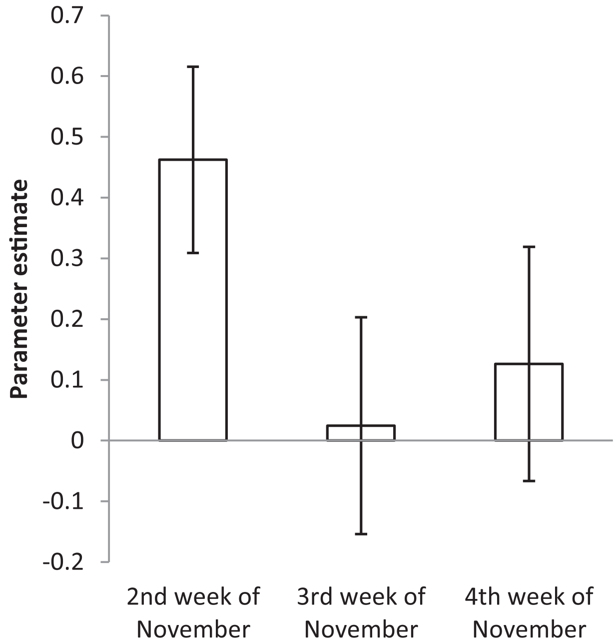

Fig. S8. Average treatment effect on the treated, where outcome variable (quantity of cross-race exchanges initiated) is measured over three distinct periods (second, third, and fourth weeks of November) corresponding to the three weeks following the treatment period (first week of November).

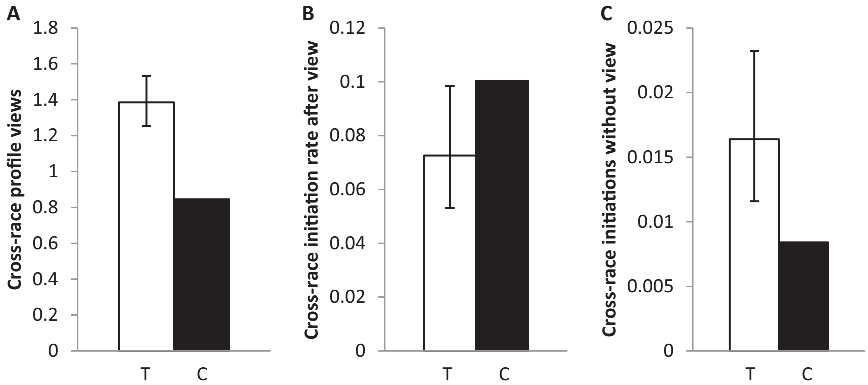

Fig. S9. Average treatment effect on the treated, where outcome variable is defined as (*A*) the quantity of cross-race profile views (negative binomial regression), (*B*) the cross-race initiation rate conditional on a profile view (binomial generalized linear model with logit link function), and (*C*) the quantity of cross-race initiations that were not preceded by a profile view (negative binomial regression).

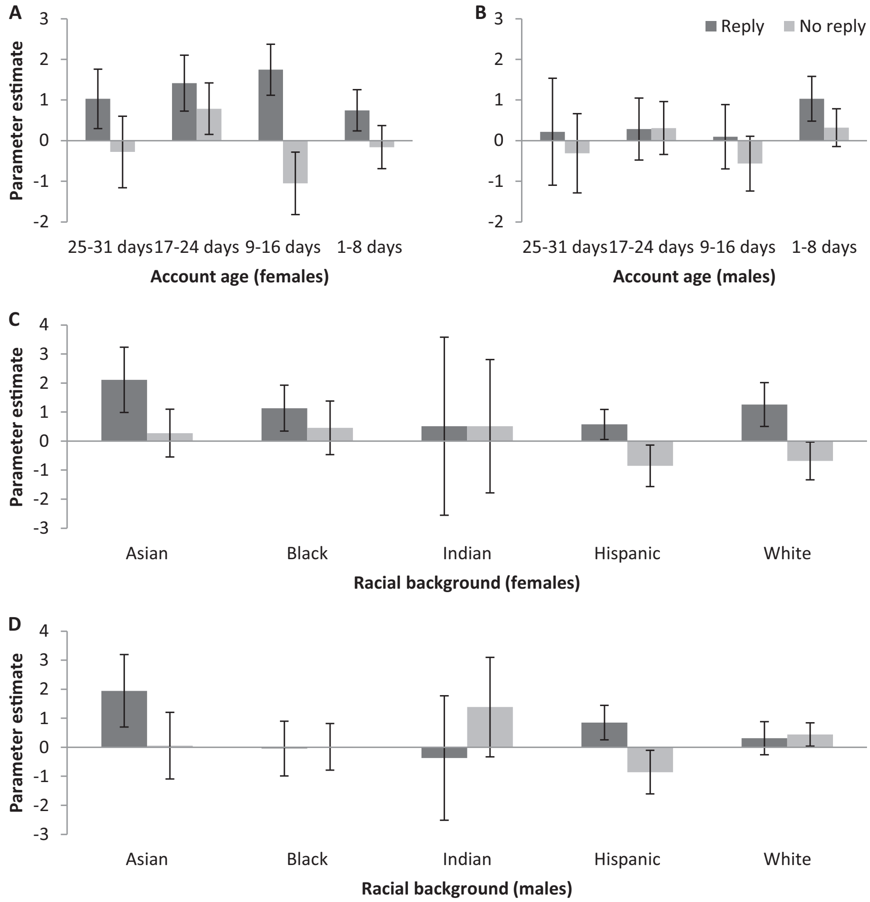

Fig. S10. Replication of Fig. 3 *B–D* from the main text, repeated separately for users who did and did not reply to the treatment message (as in Fig. 4*B*). Specifically, results are presented separately for (*A*) women with varying account ages, (*B*) men with varying account ages, (*C*) women from different racial backgrounds, and (*D*) men from different racial backgrounds.

Table S1. Demographic composition of study sample

Attribute	Male (n = 67,054)	Female (n = 59,080)
Race		
Asian	3.61	4.11
Black	5.21	6.46
Indian	0.99	0.63
Hispanic/Latin	6.64	5.90
White	83.55	82.91
Income, $		
<20,000	8.15	7.53
20,000–30,000	7.07	5.27
30,000–40,000	5.87	3.69
40,000–50,000	4.34	2.42
50,000–60,000	3.28	1.70
60,000–70,000	2.08	0.92
70,000–80,000	1.64	0.53
80,000–100,000	1.81	0.72
100,000–150,000	1.88	0.44
150,000–250,000	0.72	0.16
250,000–500,000	0.26	0.05
500,000–1,000,000	0.11	0.03
>1,000,000	0.69	0.25
Rather not say	22.72	28.28
[Unreported]	39.38	48.02
Education		
High school	13.20	10.30
2-y college	11.34	10.70
College/university	47.21	47.59
Master's program	8.95	13.80
Law school	1.43	1.53
Medical school	0.85	1.10
PhD program	2.02	1.96
Space camp	2.34	1.21
[Unreported]	12.68	11.81
Religion		
Agnosticism	10.73	8.01
Atheism	8.52	3.99
Christianity	28.18	35.64
Judaism	2.54	3.78
Catholicism	11.21	13.83
Islam	0.12	0.07
Hinduism	0.40	0.28
Buddhism	1.22	1.00
Other	10.96	10.11
[Unreported]	26.13	23.29
Age		
Mean	30.06	31.34
SD	9.97	10.82

All statistics are percentages unless otherwise indicated. Categories above are reproduced exactly as they appear on the dating site. Naturally, it is not possible to assess the veracity of these self-reported data; in at least one instance ("space camp"), the response option is clearly meant to be facetious.

Table S2. Parameter estimates from all exponential random graph models

Two-digit ZIP code region	*N*	Density	Reciprocity	Female-receiver	Asian-matching	Black-matching	Indian-matching	Hispanic-matching	White-matching	Asian-matching* reciprocity	Black-matching* reciprocity	Indian-matching* reciprocity	Hispanic-matching* reciprocity	White-matching* reciprocity
01	1,809	−8.75	6.67	1.59	0.82			1.72	0.66	84.79			−3.22	0.67
02	4,136	−8.16	7.07	1.57	1.13	1.39	2.80	0.71	0.06	−0.87	−1.53	−1.85	−0.92	0.21
06	1,646	−7.63	6.56	1.84	−0.06	0.43		0.58	−0.11		−1.14		0.07	0.45
07	2,035	−8.85	7.31	1.79	1.15	1.35	3.40	1.47	0.66	−1.38	−0.55		−0.98	−0.14
08	1,715	−8.56	7.55	1.83	0.73	1.02	1.56	1.53	0.36	0.30	−0.51	−0.34	−0.94	−0.27
10	5,599	−8.87	7.24	1.63	1.05	1.44	2.38	1.28	0.40	−0.60	−0.93	−2.15	−1.21	0.18
11	4,952	−9.02	7.55	1.65	0.88	1.21	0.34	1.13	0.48	−0.22	−0.15		−0.96	0.01
14	1,174	−8.13	7.32	2.01	2.02	1.04			0.36	−1.12	−0.62			−0.43
15	1,158	−7.51	6.36	1.64	1.62	0.95			0.30		0.37			0.04
19	2,902	−8.36	7.08	1.58	1.71	0.97	1.63	1.56	0.42	−1.15	−0.29	−0.85	−2.17	−0.06
20	2,465	−8.70	7.34	1.77	1.08	1.34	1.95	0.68	0.59	−1.13	−0.42	−1.46	0.47	−0.07
21	1,807	−8.48	7.38	1.85	1.40	0.67		0.71	0.59		0.22		0.62	−0.44
22	1,359	−8.20	6.46	1.73	1.71	0.80	2.65	1.77	0.23	−0.79	−0.30			0.57
23	1,133	−8.41	7.28	2.08	2.37	1.37		1.49	0.86		−0.48			−0.66
27	1,509	−8.41	7.33	1.45		1.39		2.11	1.03		−0.71			−0.70
28	1,658	−8.41	7.14	1.84		0.74		1.00	0.31		0.68		−0.49	0.40
29	1,115	−8.29	7.35	1.76		0.23			0.43		1.02			−0.20
30	2,632	−8.85	7.55	2.01	2.17	1.36	2.13	2.02	0.88	−2.27	−0.87		−1.31	−0.55
32	2,528	−8.49	7.37	1.78	0.47	0.65	3.27	1.02	0.21	1.14	−0.73		−1.07	0.01
33	2,872	−8.75	7.41	2.01	0.49	1.02		1.55	0.36	2.15	−0.61		−1.45	0.10
37	1,329	−7.72	6.51	1.49		0.01			0.06		1.55			0.56
43	1,337	−8.03	7.19	1.92	1.75	0.33		2.28	0.47		0.17			−0.46
44	1,762	−8.02	7.00	1.60	2.84	0.65		1.57	0.56	−1.35	0.11			−0.34
45	1,232	−7.90	7.15	1.72		1.19			0.26		−1.88			−0.37
46	1,509	−7.95	6.36	1.65		0.39		0.74	0.21		1.72			0.60
48	2,750	−8.34	7.29	1.70	2.94	1.80	2.97	0.55	0.32	−4.79	−1.35			−0.09
49	1,056	−7.52	6.59	1.57		1.11		1.17	−0.25					0.43
53	1,400	−7.87	7.16	1.48		0.14	4.18	1.35	0.37		1.36			−0.34
55	2,127	−7.61	6.55	1.62	0.88	1.14			−0.08	−0.42	−2.01			0.23
60	5,054	−8.49	7.16	1.75	1.22	1.54	2.29	1.00	0.20	−1.79	−1.10	−1.28	−0.38	0.13
75	1,780	−8.01	6.90	1.69	1.35	0.98		0.91	0.34	−1.22	−0.42		−0.07	−0.06
76	1,345	−7.82	6.55	1.73	0.64	0.63		0.32	0.11		0.38		0.52	0.33
77	2,072	−8.03	6.91	1.75	1.90	0.81		1.08	0.19	−1.89	0.12		−0.82	0.27
78	2,685	−8.14	7.10	1.57	2.14	1.07	4.13	0.54	0.40	−2.11	−1.33		−0.09	−0.16
80	2,307	−7.54	6.52	1.48	−0.14	−0.37		0.71	−0.20		1.78		−1.02	0.51
85	2,237	−8.65	7.63	2.22	1.44	0.45		0.54	0.51	−0.78	0.28		−0.38	−0.36
90	3,689	−8.57	7.12	1.67	1.00	0.83	2.23	0.72	0.41	−1.49	−0.03		−0.17	0.10
91	2,315	−9.35	8.15	2.11	1.22	0.49		1.02	0.78	−1.05	0.78		−0.59	−0.50
92	3,853	−9.12	7.84	2.03	1.21	0.66		0.87	0.59	−0.72			−0.56	−0.29
93	1,118	−8.17	7.32	2.01				0.62	−0.10				0.09	0.34
94	4,777	−8.66	7.30	1.60	1.07	0.84	1.59	0.76	0.23	−0.75	1.14	−1.60	−0.57	0.31

Table S2. Cont.

Two-digit ZIP code region	*N*	Density	Reciprocity	Female-receiver	Asian-matching	Black-matching	Indian-matching	Hispanic-matching	White-matching	Asian-matching* reciprocity	Black-matching* reciprocity	Indian-matching* reciprocity	Hispanic-matching* reciprocity	White-matching* reciprocity
95	2,564	−8.73	7.61	1.76	1.34	1.05	1.73	0.28	0.29	−1.56	−0.34		0.50	0.04
97	2,411	−8.31	7.41	1.82	1.17	0.94		1.27	0.13				−1.09	0.06
98	3,627	−8.37	7.26	1.75	0.86	1.16	1.30	−0.86	−0.11	−0.71	−0.45	1.15		0.40

I estimated a separate model for each two-digit ZIP code region with more than 1,000 users represented in the study sample. Some parameters could not be estimated because of an absence of the relevant type of tie. All models converged except for the model for ZIP code region 01xxx, as evinced by the unusually high parameter estimate for Asian same-race reciprocity; results from this model were omitted from Fig. 2.

Table S3. Numerical presentation of results visualized in the main text

Fig.	Subgroup	Constant Coef.	Constant SE	Treatment Coef.	Treatment SE	*N*
3*A*	Overall effect	−2.273***	0.035	0.317***	0.093	30,495
3*B*	Female, account age 25–31 d	−3.373***	0.126	0.572*	0.282	3,103
3*B*	Female, account age 17–24 d	−3.778***	0.135	0.971***	0.275	4,523
3*B*	Female, account age 9–16 d	−2.511***	0.125	0.071	0.291	4,479
3*B*	Female, account age 1–8 d	−3.075***	0.095	0.209	0.202	4,916
3*B*	Male, account age 25–31 d	−1.299***	0.092	−0.038	0.399	2,193
3*B*	Male, account age 17–24 d	−1.361***	0.067	0.296	0.257	3,203
3*B*	Male, account age 9–16 d	−0.997***	0.062	−0.222	0.258	3,343
3*B*	Male, account age 1–8 d	−1.398***	0.058	0.660***	0.184	4,735
3*C*	Asian female	−2.558***	0.276	1.218**	0.392	682
3*C*	Black female	−2.098***	0.180	0.723*	0.331	747
3*C*	Indian female	−2.590***	0.700	0.511	0.960	56
3*C*	Hispanic female	−1.622***	0.154	−0.302	0.255	932
3*C*	White female	−4.438***	0.093	−0.092	0.250	14,604
3*D*	Asian male	−1.424***	0.211	1.168**	0.425	190
3*D*	Black male	−0.363*	0.153	−0.014	0.316	365
3*D*	Indian male	−0.470	0.406	0.470	0.727	55
3*D*	Hispanic male	−0.237*	0.107	0.292	0.242	672
3*D*	White male	−2.092***	0.042	0.396*	0.169	12,192
4*A*	Received interracial tie previously	−2.322***	0.101	0.323	0.175	4,999
4*A*	Did not receive interracial tie previously	−2.438***	0.042	0.511***	0.127	22,256
4*B*	Replied to treatment message	−1.768***	0.036	0.765***	0.140	15,926
4*B*	Did not reply to treatment message	−2.480***	0.041	−0.079	0.122	26,873
4*C*	Specific effect	−2.905***	0.130	0.511**	0.177	6,138
4*C*	Generalized effect	−3.465***	0.141	0.099	0.197	6,138
4*D*	First week after treatment	−2.273***	0.035	0.317***	0.093	30,495
4*D*	Second week after treatment	−2.230***	0.037	0.157	0.103	30,495

All models were estimated using negative binomial regression. Results are presented here in their original logged form. In all models, treatment and control cases are matched exactly based on gender, racial background, and two-digit ZIP code region, and coarsely based on overall quantity of initiation messages sent during the control period (dichotomized as one or more messages compared with no messages), quantity of interracial initiation messages sent during the control period (dichotomized as one or more interracial messages compared with no messages), and account length (divided into four approximately equal periods based on whether users joined in the first, second, third, or fourth quarter of October). $^*P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$ (two-tailed tests).