

Is Romantic Desire Predictable? Machine Learning Applied to Initial Romantic Attraction



Samantha Joel¹, Paul W. Eastwick², and Eli J. Finkel^{3,4}

¹Department of Psychology, University of Utah; ²Department of Psychology, University of California, Davis; ³Department of Psychology, Northwestern University; and ⁴Kellogg School of Management, Northwestern University

Psychological Science
2017, Vol. 28(10) 1478–1489
© The Author(s) 2017
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0956797617714580
www.psychologicalscience.org/PS



Abstract

Matchmaking companies and theoretical perspectives on close relationships suggest that initial attraction is, to some extent, a product of two people's self-reported traits and preferences. We used machine learning to test how well such measures predict people's overall tendencies to romantically desire other people (actor variance) and to be desired by other people (partner variance), as well as people's desire for specific partners above and beyond actor and partner variance (relationship variance). In two speed-dating studies, romantically unattached individuals completed more than 100 self-report measures about traits and preferences that past researchers have identified as being relevant to mate selection. Each participant met each opposite-sex participant attending a speed-dating event for a 4-min speed date. Random forests models predicted 4% to 18% of actor variance and 7% to 27% of partner variance; crucially, however, they were unable to predict relationship variance using any combination of traits and preferences reported before the dates. These results suggest that compatibility elements of human mating are challenging to predict before two people meet.

Keywords

attraction, dating, speed dating, romantic desire, romantic relationships, machine learning, statistical learning, random forests, ensemble methods, open data, open materials

Received 9/16/16; Revision accepted 5/19/17

Achieving a high-quality romantic relationship is a goal with both evolutionary consequences (Fletcher, Simpson, Campbell, & Overall, 2015) and practical consequences (Kiecolt-Glaser & Newton, 2001). Yet the task of finding a suitable partner can be time consuming and anxiety provoking (Spielmann et al., 2013). Although identifying the most attractive person in one's social milieu might be straightforward, identifying someone who finds you *uniquely* appealing—and whom you find uniquely appealing in return—is no simple feat.

The challenges of dating have created a strong economic market for matchmaking services in which companies strive to provide their customers with tailored romantic matches. When signing up for a dating service, users complete questionnaires assessing psychological constructs that vary across individuals (e.g., values, personality, preferences for particular qualities in a partner).

The service then selects suitable potential partners by feeding the questionnaire responses into an algorithm. Many companies (e.g., chemistry.com, OKCupid.com) claim to be able to match users with partners with whom they are especially likely to “click” on first meeting. Other companies go even further, claiming that they can predict the much more distal outcome of long-term-relationship compatibility (e.g., eHarmony.com). Although these claims have not been scientifically vetted, they are not far-fetched from a theoretical point of view. Myriad perspectives in the close-relationships and evolutionary-psychology literatures suggest that outcomes such as

Corresponding Author:

Samantha Joel, Department of Psychology, University of Utah, 380 South 1530 East, Salt Lake City, UT 84112
E-mail: samantha.joel@psych.utah.edu

relationship satisfaction and longevity follow from the conjunction of two partners' preferences, traits, and personal histories (e.g., Buss & Barnes, 1986; Byrne, 1961; Campbell, Chin, & Stanton, 2016; McNulty, 2016).

For romantic-matching algorithms to be effective at all, one or more of the following conditions must be met: It must be possible, in principle, to predict (a) people's overall tendencies to romantically desire other people (actor effects), (b) people's tendencies to be desired by other people (partner effects), and (c) people's desire for specific partners above and beyond actor and partner effects (relationship effects; Eastwick & Hunt, 2014). If the first or second condition were met, an algorithm could help people form relationships by excluding people who are exceptionally misanthropic (i.e., low actor effect) or exceptionally undesirable (i.e., low partner effect)—or both—from the group of eligible daters. But the last of these components—unique desire—is the *raison d'être* behind commercial approaches to matching. That is, people are willing to pay for matching services typically because those services claim to provide matches uniquely tailored for each user that are particularly likely to lead to a relationship (Finkel, Eastwick, Karney, Reis, & Sprecher, 2012). The primary purpose of the present research was to test whether it is indeed possible to predict unique romantic desire using measures collected before the two individuals have met.

Prior Perspectives on the Predictability of Romantic Attraction

Given the current scientific knowledge base and tools, and drawing from self-report data gathered before potential partners have met, is it possible to anticipate which pairs of heterosexual individuals will be particularly interested in dating one another? A close reading of the existing empirical literature may inspire skepticism. The collected wisdom of this field has produced minimal insight into the prediction of relationship outcomes—especially outcomes measured at the level of the dyad (e.g., Partner A's feelings about Partner B)—from information collected before two people have met. As romantic relationships develop over time, couples bond over shared experiences, such as disclosing thoughts and feelings (Laurenceau, Barrett, & Pietromonaco, 1998), navigating relationship threats (Murray, Holmes, & Collins, 2006), celebrating each other's successes (Gable, Gonzaga, & Strachman, 2006), and responding to each other's needs (Impett, Gable, & Peplau, 2005). Thus, relationship success is much more than the sum or interaction of the characteristics that each person brings to the relationship (Rusbult & Van Lange, 2003).

Indeed, models such as the vulnerability-stress-adaptation model (Karney & Bradbury, 1995) and the ReCAST (relationship coordination and strategic timing) model (Eastwick, 2016) highlight the chance events, dyad-specific experiences, and chaotic forces that may cause the emergence and persistence of a relationship to be difficult or impossible to predict *a priori* (see also Eastwick, Harden, Shukusky, Morgan, & Joel, 2017; Weigel & Murray, 2000). Consistent with these models, the strongest predictors of relationship outcomes (i.e., maintenance or dissolution) tend to be features of the relationship itself—such as love, commitment, and closeness (Le, Dove, Agnew, Korn, & Mutso, 2010). These features cannot be meaningfully assessed until two people meet and begin interacting (Finkel et al., 2012) and are therefore not available to matching algorithms.

Empirical efforts to predict relationship-level variance in initial attraction from variables assessed before two people meet have also tended to fare poorly. For example, initial attraction in face-to-face contexts is negligibly related to similarity (e.g., the fact that Laura and Ben share similar interests makes them no more or less likely to be attracted to each other; Luo & Zhang, 2009; Tidwell, Eastwick, & Finkel, 2013). It is also unrelated to idiosyncratic mate preferences (e.g., the match between Laura's reported preference for extraverted men and Ben's reported extraversion makes it no more or less likely that Laura will be attracted to Ben; Eastwick, Luchies, Finkel, & Hunt, 2014). In other words, little predictive power is gained by examining which pairs of individuals share each other's traits or match each other's ideals. Many individual differences have successfully predicted people's overall tendencies to desire others and to be desired by others (e.g., McClure, Lydon, Bacus, & Baldwin, 2010; Montoya, 2008). For example, people tend to be more selective (i.e., actor variance) and more desirable (i.e., partner variance) in mating contexts to the extent that they are physically attractive (Montoya, 2008). But predicting relationship-level romantic desire—again, the primary contribution purportedly offered by any matching algorithm—may not be achievable using measures collected before the couple meets (e.g., personality, ideals, values). Rather, accurately predicting which pairs of individuals share a unique romantic connection may be possible only with the experiential, dyadic information that emerges in the wake of an initial face-to-face interaction (Finkel et al., 2012).

The Random Forests Algorithm

In the present research, we attempted to predict romantic desire as accurately as possible by taking advantage

of a method of machine learning called random forests (Breiman, 2001; Liaw & Wiener, 2015). This method is specifically designed to answer questions about prediction and holds two key advantages over conventional regression models (Strobl, Malley, & Tutz, 2009). First, random forests can handle many predictors at once while minimizing overfitting. Second, random forests are sensitive to nonlinear relationships, including complex interactions among predictors. In essence, random forests allow us to (a) simultaneously test a wide range of psychological measures that may predict romantic desire, rather than only a subset, and (b) account for all potential interactions between two people's responses that might contribute to their unique desire for each other. Thus, this study aimed to provide the most thorough and comprehensive test to date of the notion that romantic attraction can be predicted from self-reported traits and preferences.

Method

In two samples of speed daters, we used random forests (Liaw & Wiener, 2015) to predict romantic desire. As described briefly earlier, random forests are a technique for machine learning that can identify robust predictors of an outcome. The two major advantages of machine learning are as follows. First, with conventional regression, all predictors work in concert to predict all dependent observations. Regression can thus accommodate only as many predictors as there are observations, and overfitting and collinearity become issues of increasing concern as more predictors are added to the model. Random forests, on the other hand, bootstrap subsamples of predictors and observations, which gives each predictor opportunities to contribute to the model without competing against more dominant predictors. This method can thus handle many predictors—even more predictors than there are observations—while remaining relatively robust against problems of overfitting and collinearity.

A second key advantage of random forests is that they are nonparametric; that is, they do not impose a particular structure to the data. As such, random forests can identify potentially complex interactions among predictors. Such interactions might be intuitive (e.g., a partner's extraversion is a strong predictor of an actor's romantic desire, particularly for actors who say that they want extraverted partners; Eastwick et al., 2014) or nonintuitive (e.g., a partner's extraversion is a strong predictor of an actor's romantic desire, particularly for actors who have low self-esteem) given existing theory. Whereas a conventional regression model cannot account for such interactions unless specified by the researcher, random forests can and will detect such

interactions, provided that the interactions meaningfully contribute to the model's overall predictive power.

Participants

Sample A consisted of 163 undergraduate students (81 women and 82 men; mean age = 19.6 years, $SD = 1.0$) who attended one of seven speed-dating events in 2005. Sample B consisted of 187 undergraduate students (93 women and 94 men; mean age = 19.6 years, $SD = 1.2$) who attended one of eight such events in 2007. Sample size was determined by the number of speed-dating events we were able to hold in 2005 and 2007 and the number of participants we were able to recruit for each event while maintaining an equal gender ratio. All participants, who were recruited via on-campus flyers and e-mails to participate in a speed-dating study, had the goal of meeting and potentially matching with opposite-sex participants. Detailed descriptions of the speed-dating research procedures and characteristics of each sample can be found in two previously published papers (Finkel, Eastwick, & Matthews, 2007; Tidwell et al., 2013).

Materials and procedure

Predictors. Participants first completed a 30-min online questionnaire that included a wide range of psychological constructs, including personality measures (e.g., the Big Five personality dimensions, attachment style, perceptions of one's own mate value), well-being assessments (e.g., positive affectivity, negative affectivity, satisfaction with life), mating strategies (e.g., sociosexuality, interest in long-term relationships), values (e.g., traditionalism, conservatism), and self-reported traits (e.g., warmth, physical attractiveness) along with ideal-partner-preference items for those same traits.

Broadly speaking, we used two procedures for generating the measures on this questionnaire. First, we culled a large set of constructs that are commonly used in major studies in the relationships literature. The starting point for this process was a set of longitudinal studies spearheaded by leading relationship scientist Caryl Rusbult in the late 1990s and early 2000s (Kumashiro, Finkel, & Rusbult, 2002). Eli Finkel, a coauthor on the present study and a former student of Rusbult's, adopted or adapted these measures—and added a handful of new ones—for a study of first-year college students in 2003–2004 (Finkel, Burnette, & Scissors, 2007). When making decisions about which measures to include in the current study, we relied heavily on that Finkel study. Second, we reviewed the social psychological literature on attraction and the evolutionary psychological literature on human mating,

incorporating several individual-differences constructs from those literatures as well.

The full 30-min questionnaire was designed to be maximally comprehensive of these fields; indeed, the constructs we prioritized are widely used (collectively cited 96,236 times as of March 1, 2017; for references, see Databases S1 and S2 in the Supplemental Material available online) and are predictive of attraction and relationship-relevant outcomes (e.g., neuroticism, Karney & Bradbury, 1995; attachment style, Kirkpatrick & Davis, 1994; sociosexuality, Simpson & Gangestad, 1991; approach/avoidance goals, Gable & Impett, 2012; warmth-trustworthiness, vitality-attractiveness, and status-resources traits, Fletcher, Simpson, Thomas, & Giles, 1999).

For the key analyses in the present article, we included nearly all the psychological constructs as predictors (182 constructs in Sample A and 112 constructs in Sample B). We omitted highly exploratory items (e.g., “What are your three favorite television shows?”), as well as several items with unusual response scales (e.g., “Do you expect that your future spouse will work full-time, part-time, or not at all if/when you have young children (i.e., before they start school)?”).¹

Of the items included in present analyses, 8% of Sample A items and 19% of Sample B items were also included in analyses reported in articles published previously (see Databases S1 and S2 in the Supplemental Material; means, standard deviations, and ranges are also provided for each continuous measure). Variability was generally substantial across these measures: Across samples, most continuous variables had a standard deviation of at least 1 (87% for Sample A, 83% for Sample B), and a range of at least 5 on either a 7-point scale (88% of 76 measures in Sample A; 88% of 57 measures in Sample B) or a 9-point scale (89% of 100 measures in Sample A; 67% of 41 measures in Sample B).² Thus, there is little reason to expect that these variables would collectively fail to predict romantic desire a priori on the basis of insufficient variability (cf. Li et al., 2013).

Approximately 1 to 2 weeks after completing the intake questionnaire, participants attended a speed-dating event in which they had a series of 4-min speed dates with approximately 12 members of the opposite sex. Immediately after each speed date, participants filled out a 2-min interaction record questionnaire containing items that assessed their experiences on their most recent speed date. In subsidiary analyses (see the Subsidiary Random Forests Analyses section), we used most of these constructs (18 in Sample A, 20 in Sample B) as predictors in the random forests models (for all postinteraction measures, see Databases S3 and S4 in the Supplemental Material).

Dependent measure. On the interaction-record questionnaire, participants completed a three-item measure of their romantic desire for that individual: “I really liked my interaction partner,” “I was sexually attracted to my interaction partner,” and “I am likely to say ‘yes’ to my interaction partner.” These items were rated on a 9-point scale (1 = *strongly disagree*, 9 = *strongly agree*). For Sample A, α was .88 ($M = 5.04$, $SD = 2.11$); for Sample B, α was .87 ($M = 4.93$, $SD = 1.90$).

Results

Sources of variance

It was essential to first confirm that our dependent measure—romantic desire reported in the wake of a 4-min interaction—comprises actor variance (how much participants desired their speed-dating partners on average), partner variance (how much participants were desired by their speed-dating partners on average), and relationship variance (how much participants desired particular partners above and beyond the participants’ actor effects and the partners’ partner effects). If any of these variances were zero or near zero, then it would not be possible to predict that source of variance from any conceivable collection of predictors.

We therefore conducted a series of social-relations-model analyses (using the BLOCKO program; Kenny, 1998) in which romantic desire was partitioned into actor, partner, and relationship variance. These analyses revealed that a nontrivial percentage of romantic desire in the present samples could be attributed to each of these three sources (Table 1). Relationship variance was the largest source of variance, followed by partner variance and then actor variance; all three exceeded the “meaningful” threshold of 10% (Kenny, Kashy, & Cook, 2006). In other words, these studies were ideal for testing questions about the ability to predict actor, partner, and relationship variance, because all three were present in the dependent measure. (If anything, it might be easiest to predict relationship variance, given that it was the largest source of variance.)

We next separated each report of romantic desire (e.g., Male 1’s reported desire for each of his 12 speed dates) into these three statistically independent components. First, we calculated actor desire—the extent to which the participant liked his or her speed-dating partners on average—by subtracting the romantic desire grand mean from the average of each participant’s approximately 12 reports of romantic desire. Second, we calculated partner desire—the extent to which the participant was liked by his or her speed-dating partners on average—by subtracting the romantic desire

Table 1. Results From Analyses With the Social-Relations Model

| Sample and statistic | Actor desire | Partner desire | Relationship desire | | Error | |
|----------------------|--------------|----------------|------------------------|------------------------|-------|-------|
| | | | Men's desire for women | Women's desire for men | Men | Women |
| Sample A | | | | | | |
| Variance | 12.15% | 25.90% | 34.74% | 31.0% | 27.3% | 31.0% |
| Reliability | .71 | .88 | .85 | .84 | | |
| Sample B | | | | | | |
| Variance | 13.60% | 22.52% | 35.98% | 32.1% | 27.9% | 31.8% |
| Reliability | .78 | .86 | .85 | .82 | | |

grand mean from the average of the approximately 12 reports of romantic desire about that participant. Third, we calculated relationship desire—the extent to which the participant liked a particular partner above and beyond his or her actor effect and the partner's partner effect—by subtracting the grand mean, the participant's actor effect, and the partner's partner effect from the participant's report of romantic desire for that partner. In our analyses, we attempted to predict each of these three components separately.

Strategy for random forests analysis

For models predicting actor and partner desire, data sets were organized at Level 2, such that each participant was represented by a row. Thus, actor and partner analyses for Sample A had 182 predictors and 163 rows, and actor and partner analyses for Sample B had 112 predictors and 187 rows. Gender was included as a predictor for these analyses (and, as part of the random-forests algorithm, as a potential moderator of any other possible effect).

For models predicting relationship desire, data sets were organized at Level 1, such that each observation was a dyad. Thus, each participant was represented approximately 12 times: once for each of their dates. Each predictor variable was included twice: once representing the value for the male member of the dyad (e.g., his extraversion), and once representing the value for the female member of the dyad (e.g., her extraversion). We conducted separate analyses predicting men's unique desire for women and women's unique desire for men. Overall, relationship analyses for Sample A had 362 predictors (i.e., 181 Sample A predictors for the man in the dyad and 181 predictors for the woman) and 958 rows, and relationship analyses for Sample B had 222 predictors and 1,092 rows. Normally, multilevel methods would allow a data analyst to enter each dyad twice—representing each member of the dyad as both an actor and a partner—such that men's desire for

women and women's desire for men could be tested together in a single analysis (Kenny et al., 2006). However, such techniques have not yet been developed for use with random forests. Thus, we tested men and women separately to avoid violating independence assumptions. As the results reveal, the (negligible) effects were comparable for men and women.

The data were analyzed using the *randomForest* package (Liaw & Wiener, 2015) for the R software environment (R Development Core Team, 2016). For all analyses, we set “ntree” to 5,000, which means that each model was constructed from 5,000 regression trees, and we left “mtry”—the number of predictors available for splitting at each tree node—at its default value of one third of the total number of predictors. For each model, we report the mean squared error (*MSE*) and the percentage of variance explained for each model, both of which the algorithm calculates using out-of-bag (OOB) observations.

Variable selection was conducted using the *VSURF* package for R (Genuer, Poggi, & Tuleau-Malot, 2010, 2016). We constructed models using variable selection criteria at three levels of stringency. The threshold step of *VSURF* eliminated variables that failed to reduce the model's error rate (liberal selection). The interpretation step of *VSURF* eliminated variables that failed to reduce the model's error rate by a sufficient amount, as determined by *VSURF*'s statistical cutoffs (moderate selection). Finally, the prediction step of *VSURF* minimized the number of predictors but maintained predictive power (stringent selection). (For procedural details on how *VSURF* selects predictors, see Genuer et al., 2010, 2016.)

We also constructed models in which no selection criteria were used, such that all predictors were included in each model (see Table S1 in the Supplemental Material). The amount of variance explained was substantially worse without the use of variable selection, which suggests that including many irrelevant predictors harmed the models' predictive power. However, the

Table 2. Summary of Results From the Primary Random Forests Models Predicting Actor, Partner, and Relationship Desire in Samples A and B

| Dependent measure and variable selection ^a | Sample A | | | Sample B | | |
|---|----------------------|------|--------------------------|----------------------|------|--------------------------|
| | Number of predictors | MSE | Total variance explained | Number of predictors | MSE | Total variance explained |
| Actor desire | | | | | | |
| Liberal | 41 | 0.84 | 17.78% | 44 | 0.84 | 4.95% |
| Moderate | 17 | 0.88 | 15.88% | 14 | 0.81 | 8.25% |
| Stringent | 9 | 0.88 | 15.42% | 5 | 0.80 | 9.52% |
| Partner desire | | | | | | |
| Liberal | 59 | 1.35 | 19.70% | 47 | 0.97 | 24.64% |
| Moderate | 12 | 1.32 | 21.43% | 7 | 0.94 | 26.70% |
| Stringent | 9 | 1.30 | 22.14% | 2 | 1.05 | 18.48% |
| Relationship desire (men for women) | | | | | | |
| Liberal | 16 | 1.90 | -4.55% | 52 | 1.70 | -3.10% |
| Moderate | 1 | 1.82 | -0.18% | 2 | 1.67 | -1.42% |
| Stringent | 1 | 1.82 | -0.18% | 2 | 1.67 | -1.42% |
| Relationship desire (women for men) | | | | | | |
| Liberal | 39 | 2.09 | -1.65% | 1 | 1.77 | -2.68 |
| Moderate | 20 | 2.07 | -1.03% | 1 | 1.77 | -2.68 |
| Stringent | 3 | 2.02 | 1.34% | 0 | — | — |

Note: *Actor desire* refers to how desirable a participant found his or her partners to be. *Partner desire* refers to how desirable a participant's partners found him or her to be. *Relationship desire* refers to desire for a particular partner, beyond actor effects and partner effects.

^aThe table shows results for each dependent variable and sample. We ran three models in which different numbers of predictors were included. In the models with liberal variable selection, we eliminated only irrelevant variables; in the model with moderate variable selection, we kept moderately predictive variables; and in the stringent model, we kept only the most predictive variables.

amount of variance explained by the models varied little regardless of which selection criterion was used (see Table 2).

Random forests results

Overall, the key random forests analyses drew from 181 traits and preferences in Sample A and 112 traits and preferences in Sample B to predict four dependent variables in each sample: general tendency to desire others (actor desire), general tendency to be desired (partner desire), men's particular desire for each woman (male relationship desire), and women's particular desire for each man (female relationship desire). Results can be seen in Table 2.

The resulting models predicted approximately 5% to 18% of the variance in actor desire and 18% to 27% of the variance in partner desire. That is, random forests could account for a modest amount of the variance in how much people tended to desire, and be desired by, their speed-dating partners in general. Consistent predictors of actor desire (i.e., the tendency to desire others) included desired level of warmth and responsiveness

in a speed date and one's own expected selectivity when choosing dates (see Tables S2 and S3 in the Supplemental Material). In other words, people who see warmth as an attractive quality tended to experience greater attraction for their dates on average, and people who expected to be more selective tended to experience less attraction for their dates on average. Consistent predictors of partner desire (i.e., the tendency to be desired by others) included participants' self-reports of their own mate value and physical attractiveness (see Tables S4 and S5 in the Supplemental Material). These results suggest that people have knowledge of their own attractiveness; people with self-reported high mate value and high physical attractiveness were indeed more desired by their dates.

In contrast, models predicted between -4.55% and -0.18% of variance in men's desire for women, and between -2.68% and 1.30% of variance in women's desire for men. (Predictors selected in each model are presented in Tables S6 through S9 in the Supplemental Material.) Furthermore, predictors were not consistent across models; indeed, many of the relationship models explained a negative percentage of variance. The

Table 3. Summary of Random Forests Models Trained on Sample A and Tested on Sample B

| Dependent measure | Number of predictors | Sample A MSE | Variance explained in Sample A | Test MSE (Sample B) | Correlation between predicted and actual scores (Sample B) |
|-------------------------------------|----------------------|--------------|--------------------------------|---------------------|--|
| Actor desire | 10 | 0.88 | 15.48% | 0.88 | .19* |
| Partner desire | 13 | 1.32 | 21.49% | 1.26 | .26** |
| Relationship desire (men for women) | 1 | 1.86 | -2.19% | 1.68 | -.06 |
| Relationship desire (women for men) | 1 | 2.07 | -0.56% | 1.76 | .02 |

Note: *Actor desire* refers to participants' responses. *Partner desire* refers to how desirable a participant's partners found him or her to be. *Relationship desire* refers to desire for a particular partner, beyond actor effects and partner effects.

* $p < .01$. ** $p < .001$.

percentage of variance explained is computed by the *randomForest* package as

$$1 - \frac{MSE_{OOB}}{\hat{\sigma}_y^2} \times 100,$$

where MSE_{OOB} is the model's mean squared OOB error, and $\hat{\sigma}_y^2$ is the variance of the dependent observations (Liaw & Wiener, 2015). Therefore, a negative score for percentage of variance explained means that the model's mean squared error is higher than the amount of variance in the dependent measure. In the context of the present data, negative variance means that the model can predict attraction less accurately than simply predicting the grand mean for every pairing. In sum, random forests were generally unable to account for *any* of the variance in how much men and women especially desired each of their matches, beyond their global tendencies to desire (actor variance) and to be desired (partner variance).

Training and testing analyses

An advantage of machine learning procedures such as random forests is that models that have been trained on one data set can then be used to predict outcome measures in different data set. Thus, these techniques are designed to answer questions about prediction in a truly a priori way. We next constructed additional models using data only from Sample A (the training data) and considering only the 87 predictors that were available in both data sets (for shared variables, see the "Shared Across Samples" column in Databases S1 and S2 in the Supplemental Material). Variables were selected for each model using the interpretation step of the *VSURF* package (moderate variable selection). We applied the training models to the equivalent predictors in Sample B, which allowed us to generate predicted scores for actor, partner, men's relationship, and

women's relationship desire in Sample B. We then compared our generated desire scores with Sample B's actual desire scores to determine how well we were truly able to predict these dependent variables (see Table 3).

The predicted actor-desire scores for Sample B correlated positively with the actual actor-desire scores for Sample B, $r = .19$, 95% confidence interval (CI) = [.05, .33], and the predicted partner-desire scores correlated positively with the actual partner-desire scores, $r = .26$, 95% CI = [.12, .39]. In contrast, men's predicted relationship-desire scores for Sample B, if anything, correlated negatively with men's actual relationship-desire scores, $r = -.06$, 95% CI = [-.121, -.002], and women's predicted relationship-desire scores for Sample B did not correlate with women's actual relationship-desire scores, $r = .02$, 95% CI = [-.04, .08]. At best, we could predict less than 0.1% of the variance in relationship desire in Sample B using the random forests models developed with Sample A. Conceptually, this means that if we know how people rate themselves on a variety of mating-relevant variables, we can use the models developed with Sample A to anticipate, with some degree of accuracy, how much they will tend to desire other people and how desirable they will be to other people in a speed-dating context. However, we cannot anticipate how much those individuals will uniquely desire each other in a speed-dating context with any meaningful level of accuracy. (Selected predictors in each final, trained model are presented in Table S10 in the Supplemental Material.)

Subsidiary random forests analyses

The random forests algorithm is relatively new to the social sciences and has rarely been applied to dyadic data. Therefore, one potential explanation for the current findings is that random forests are simply unable to capture meaningful amounts of variance in relationship desire. To address this possibility, we next conducted

Table 4. Random Forests Models Predicting Relationship Desire in Samples A and B Using Postinteraction Predictors

| Dependent measure and variable selection ^a | Sample A | | | Sample B | | |
|---|----------------------|------|--------------------------|----------------------|------|--------------------------|
| | Number of predictors | MSE | Total variance explained | Number of predictors | MSE | Total variance explained |
| Relationship desire (men for women) | | | | | | |
| Liberal | 36 | 1.35 | 26.33% | 40 | 1.17 | 28.70% |
| Moderate | 19 | 1.35 | 26.22% | 12 | 1.15 | 29.26% |
| Stringent | 5 | 1.42 | 21.67% | 1 | 1.32 | 19.67% |
| Relationship desire (women for men) | | | | | | |
| Liberal | 35 | 1.59 | 23.47% | 40 | 1.22 | 27.12% |
| Moderate | 19 | 1.58 | 24.00% | 19 | 1.27 | 26.09% |
| Stringent | 1 | 1.72 | 16.54% | 2 | 1.36 | 20.97% |

Note: *Relationship desire* refers to desire for a particular partner, beyond actor effects and partner effects.

^aThe table shows results for each dependent variable and sample. We ran a model in which different numbers of predictors were included. In the model with liberal variable selection, we eliminated only irrelevant variables; in the model with moderate variable selection, we kept moderately predictive variables; and in the stringent model, we kept only the most predictive variables.

additional analyses in which measures from the interaction record questionnaire (i.e., those completed after each speed date, alongside the dependent measure) were entered as predictors. Whereas the background questionnaire items used in our initial analyses are about the individual (i.e., each person's traits and preferences), these postinteraction measures are about perceptions of each date. These analyses test whether Partner A's particular desire for Partner B—over and above Partner A's tendencies to desire and Partner B's tendencies to be desired—can be predicted by each partner's perception of the quality of the interaction they shared with each other.

For Sample A, the predictors were 18 postinteraction measures that participants completed after each speed date (e.g., perceived chemistry with the date, perceived intelligence of the date). For Sample B, the predictors were 20 postinteraction measures. In both samples, the only postinteraction measures omitted as predictors were the three-item measure of romantic desire (i.e., the dependent measure) and the item "I knew this person very well before today's event." Analyses were conducted at Level 1. In total, Sample A included 36 predictors (18 male and 18 female predictors) and 958 rows, and Sample B included 38 predictors (20 male and 20 female predictors) and 1,092 rows. Separate analyses predicting male and female relationship desire were conducted, using the same analysis strategy used for the primary random forests models reported above.

Results are presented in Table 4. Unlike the original models, which were constructed with background

questionnaire measures, these models constructed with postinteraction measures predicted approximately 21% to 29% of male relationship desire and 16% to 24% of female relationship desire. The best predictor across both samples and both sexes was feelings of chemistry with a partner (see Tables S11–S14 in the Supplemental Material). Thus, it is not the case that desire for a specific partner could not be predicted in principle. Rather, desire for a specific partner could not be predicted from traits and preferences measured before the dyad had met.

For the sake of completeness, we also tested models in which interaction record questionnaire measures organized at Level 2 were used to predict actor and partner desire. Each person's gender, their average perceptions of their dates on each interaction-record construct, and their dates' average perceptions of them on each interaction-record construct were entered as predictors in each model. Sample A models included 37 predictors and 163 rows, and Sample B models included 41 predictors and 187 rows (for full results, see Table S15 in the Supplemental Material). People's postinteraction perceptions of their dating experiences were highly effective at predicting actor desire (72%–83% of variance explained) and partner desire (92%–94% of variance explained). The consistent predictors of actor desire were the participants' judgment of the dates' physical attractiveness and the participants' feelings of chemistry on their dates. The most consistent predictors of partner desire were the partners' judgment of the participants' physical attractiveness and the partners' feelings of chemistry with the participants.

As an alternative way to test the validity of the random forests method, we generated speed-dating data sets in which the romantic-desire dependent variable was constructed from a combination of randomly generated actor effects (i.e., simulated preferences), randomly generated partner effects (i.e., simulated traits), or randomly generated relationship effects (i.e., simulated interactions of preferences and traits). In other words, in some of these data sets, romantic desire was a function of interactions between characteristics of the two partners, and in other data sets, romantic desire was a function only of characteristics of the actor or of the partner (or both). Then, we used random forests to identify whether there were robust predictors of actor, partner, and relationship desire (just as in the primary random forests models) from among the predictors that were actually used to create the desire dependent variable (see Table S16 in the Supplemental Material).

As expected, random forests were able to predict actor desire, partner desire, and relationship desire to the extent that the dependent variable originally comprised preferences, traits, and interactions of preferences and traits, respectively. Furthermore, the accuracy levels of the random forests models were akin to those obtained with linear regression, even though preference-by-trait interactions were prespecified in the regression models but not in the random forests models. The models also performed well when additional randomly generated variables were included among the possible predictors (see Table S17 in the Supplemental Material). These results suggest that if interaction effects exist, random forests are capable of detecting them. Thus, the inability of random forests to predict relationship variance in Table 2 is consistent with the possibility that relationship desire does not consist of predictable interaction effects among background variables in our speed-dating data.

Discussion

In the present research, we sought to predict initial romantic desire as accurately as possible across two speed-dating studies, using machine learning and more than 100 self-report measures collected before a set of speed-dating events. We found that random forests could predict 4% to 18% of the variance in actor desire and 7% to 27% of the variance in partner desire. These results suggest that relationship science has uncovered many traits and preferences that can meaningfully predict people's tendencies to desire others (e.g., pickiness, self-assessments of warmth; see Tables S2 and S3 in the Supplemental Material) and to be desired by others (e.g., sociosexuality, mate value; see Tables S4 and S5 in the Supplemental Material). However, models

drawing from these background measures were consistently unable to predict any variance in relationship desire, or how much one person especially desired another person. Relationship desire was not predictable from background measures even though (a) it accounted for the largest percentage of variance in the dependent measure (Table 1) and (b) random forests were successful at predicting relationship desire using relationship-specific postinteraction measures (16%–29% of variance explained).

One possible interpretation of these findings is that it would be possible to collect background measures that would predict relationship-level desire if relationship science were to reveal what these measures might be. A related possibility is that relationship desire consists of a great many actor-by-partner interaction effects—each one tiny but real. To achieve confidence in any given interaction effect, relationship scientists might need to take lessons from the genomics literature (Hewitt, 2012; Okbay et al., 2016) and incorporate very large samples (i.e., many thousands of participants) and stringent corrections for multiple comparisons. Alternatively, relationship desire may simply not be predictable from measures collected before two individuals meet—romantic attraction may emerge from dyad-specific factors that cannot be anticipated a priori. This interpretation is supported by research highlighting the role of the situation in relationship development (Rusbult & Van Lange, 2003), as well as by mounting evidence that the romantic standards that people hold in the abstract only weakly match the qualities that they initially find romantically desirable (Eastwick et al., 2014) and the initial mate choices that they make (Joel, Teper, & MacDonald, 2014).

The present findings address only obliquely the predictability of long-term romantic compatibility. Even if unique desire in initial interactions is not predictable a priori, a matching algorithm could serve a useful function by surrounding users with partners with whom they would ultimately enjoy long-term compatibility should a relationship develop. Building and validating such an algorithm would require that researchers collect background measures before two partners have met and follow them over time as they become an established couple. To our knowledge, relationship science has yet to accomplish this methodological feat; even the commonly assessed individual-difference predictors of relationship satisfaction and breakup (e.g., neuroticism, attachment insecurity; Karney & Bradbury, 1995; Le et al., 2010) have never been assessed before the formation of a relationship. For these variables to be useful in a long-term compatibility algorithm that also separates actor, partner, and relationship variance, researchers would need to predict relationship dynamics across

participants' multiple romantic relationships over time (Eastwick et al., 2017). Predicting long-term compatibility may be more challenging than predicting initial romantic desire.

The present results were obtained with undergraduate samples; a more demographically diverse sample might exhibit matching by sociological factors such as age, socioeconomic status, cultural background, or religious background. That is, relationship effects might be predictable in a getting-acquainted context that involves greater demographic diversity (e.g., 20-year-olds are likely to prefer dating 20-year-olds to dating 40-year-olds, and 40-year-olds are likely to prefer dating 40-year-olds to dating 20-year-olds). Further, the present study examined romantic desire experienced after a 4-min interaction. It is unclear if individual characteristics (assessed before an initial interaction) would become more predictive of relationship desire over time, or if only features of the relationship itself predict relationship desire as people become better acquainted.

Initial romantic desire is a virtual prerequisite to long-term relationship success, at least in modern Western culture; two people must first like each other enough to decide to spend more time together. The goal of the present research was to test the basic assumption that initial romantic desire is predictable. Is romantic desire like a chemical reaction, such that the right combination of traits and preferences from two people will predictably result in strong levels of desire? Or, is it more like an earthquake, such that the dynamic and chaos-like processes that cause its occurrence require considerable additional scientific inquiry before prediction is realistic (Silver, 2012)? The current study suggests that the latter may be more likely than the former: Relationship desire could not be predicted from initial conditions despite the use of cutting-edge statistical methods and a vast catalog of psychological variables that have been widely cited in the field of relationship science. These findings highlight how the science of romantic relationships has much to learn from other prediction sciences before scholars can fully address this vexing and timeless question.

Action Editor

Steven W. Gangestad served as action editor for this article.

Author Contributions

Study design and data collection were completed by P. W. Eastwick and E. J. Finkel. S. Joel and P. W. Eastwick developed the research questions and the hypotheses. S. Joel conducted the analyses. S. Joel and P. W. Eastwick interpreted the results. S. Joel drafted the manuscript, and P. W. Eastwick and E. J. Finkel provided critical revisions. All the authors approved the final version of the manuscript for submission.

Acknowledgments

We thank Jacob Westfall and Tal Yarkoni for providing feedback during the development of this work.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

This research was supported by a postdoctoral fellowship from the Social Sciences and Humanities Research Council of Canada (to S. Joel).

Supplemental Material

Additional supporting information can be found at <http://journals.sagepub.com/doi/suppl/10.1177/0956797617714580>

Open Practices



All data and materials (including R syntax for all analyses) have been made publicly available via the UK Data Service (<http://reshare.ukdataservice.ac.uk/852716/>) and the Open Science Framework (<https://osf.io/4d3b9/>), respectively. The complete Open Practices Disclosure for this article can be found at <http://journals.sagepub.com/doi/suppl/10.1177/0956797617714580>. This article has received badges for Open Data and Open Materials. More information about the Open Practices badges can be found at <https://www.psychologicalscience.org/publications/badges>.

Notes

1. For all measures collected at intake, see Databases S1 and S2 in the Supplemental Material. Items included as predictors in the present analyses are listed in Sheet 1, and items not included are listed in Sheet 2, in the order in which the responses to the items were collected from the participants.
2. The range of scores was 5 points across measures, on either 7-point or 9-point scales (e.g., scores might range from 1 to 5, from 2 to 6, from 3 to 7, from 4 to 8, or from 5 to 9).

References

- Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32.
- Buss, D. M., & Barnes, M. (1986). Preferences in human mate selection. *Journal of Personality and Social Psychology*, 50, 559–570.
- Byrne, D. (1961). Interpersonal attraction and attitude similarity. *The Journal of Abnormal and Social Psychology*, 62, 713–715.
- Campbell, L., Chin, K., & Stanton, S. C. E. (2016). Initial evidence that individuals form new relationships with partners that more closely match their ideal preferences. *Collabra*, 2, 2–7.

- Eastwick, P. W. (2016). The emerging integration of close relationships research and evolutionary psychology. *Current Directions in Psychological Science, 25*, 183–190.
- Eastwick, P. W., Harden, K. P., Shukusky, J. A., Morgan, T. A., & Joel, S. (2017). Consistency and inconsistency among romantic partners over time. *Journal of Personality and Social Psychology, 112*, 838–885.
- Eastwick, P. W., & Hunt, L. L. (2014). Relational mate value: Consensus and uniqueness in romantic evaluations. *Journal of Personality and Social Psychology, 106*, 728–751.
- Eastwick, P. W., Luchies, L. B., Finkel, E. J., & Hunt, L. L. (2014). The predictive validity of ideal partner preferences: A review and meta-analysis. *Psychological Bulletin, 140*, 623–665.
- Finkel, E. J., Burnette, J. L., & Scissors, L. E. (2007). Vengefully ever after: Destiny beliefs, state attachment anxiety, and forgiveness. *Journal of Personality and Social Psychology, 92*, 871–886.
- Finkel, E. J., Eastwick, P. W., Karney, B. R., Reis, H. T., & Sprecher, S. (2012). Online dating: A critical analysis from the perspective of psychological science. *Psychological Science in the Public Interest, 13*, 3–66.
- Finkel, E. J., Eastwick, P. W., & Matthews, J. (2007). Speed-dating as an invaluable tool for studying romantic attraction: A methodological primer. *Personal Relationships, 14*, 149–166.
- Fletcher, G. J. O., Simpson, J. A., Campbell, L., & Overall, N. C. (2015). Pair-bonding, romantic love, and evolution: The curious case of *Homo sapiens*. *Perspectives on Psychological Science, 10*, 20–36.
- Fletcher, G. J. O., Simpson, J. A., Thomas, G., & Giles, L. (1999). Ideals in intimate relationships. *Journal of Personality and Social Psychology, 76*, 72–89.
- Gable, S. L., Gonzaga, G. C., & Strachman, A. (2006). Will you be there for me when things go right? Supportive responses to positive event disclosures. *Journal of Personality and Social Psychology, 91*, 904–917.
- Gable, S. L., & Impett, E. A. (2012). Approach and avoidance motives and close relationships. *Social & Personality Psychology Compass, 6*, 95–108.
- Genuer, R., Poggi, J. M., & Tuleau-Malot, C. (2010). Variable selection using random forests. *Pattern Recognition Letters, 31*, 2225–2236.
- Genuer, R., Poggi, J. M., & Tuleau-Malot, C. (2016). VSURF: Variable selection using random forests (Version 1.0.3) [Computer software]. Retrieved from <https://cran.r-project.org/package=VSURF>
- Hewitt, J. K. (2012). Editorial policy on candidate gene association and candidate gene-by-environment interaction studies of complex traits. *Behavior Genetics, 42*, 1–2.
- Impett, E. A., Gable, S. L., & Peplau, L. A. (2005). Giving up and giving in: The costs and benefits of daily sacrifice in intimate relationships. *Journal of Personality and Social Psychology, 89*, 327–344.
- Joel, S., Teper, R., & MacDonald, G. (2014). People overestimate their willingness to reject potential romantic partners by overlooking their concern for others. *Psychological Science, 25*, 2233–2240.
- Karney, B. R., & Bradbury, T. N. (1995). The longitudinal course of marital quality and stability: A review of theory, methods, and research. *Psychological Bulletin, 118*, 3–34.
- Kenny, D. A. (1998). BLOCKO [Computer software]. Available from <http://davidakenny.net/srm/srmp.htm>
- Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic data analysis*. New York, NY: Guilford Press.
- Kiecolt-Glaser, J. K., & Newton, T. L. (2001). Marriage and health: His and hers. *Psychological Bulletin, 127*, 472–503.
- Kirkpatrick, L. A., & Davis, K. E. (1994). Attachment style, gender, and relationship stability: A longitudinal analysis. *Journal of Personality and Social Psychology, 66*, 502–512.
- Kumashiro, M., Finkel, E. J., & Rusbult, C. E. (2002). Self-respect and pro-relationship behavior in marital relationships. *Journal of Personality, 70*, 1009–1049.
- Laurenceau, J. P., Barrett, L. F., & Pietromonaco, P. R. (1998). Intimacy as an interpersonal process: The importance of self-disclosure, partner disclosure, and perceived partner responsiveness in interpersonal exchanges. *Journal of Personality and Social Psychology, 74*, 1238–1251.
- Le, B., Dove, N. L., Agnew, C. R., Korn, M. S., & Mutso, A. A. (2010). Predicting nonmarital romantic relationship dissolution: A meta-analytic synthesis. *Personal Relationships, 17*, 377–390.
- Li, N. P., Yong, J. C., Tov, W., Sng, O., Fletcher, G. J. O., Valentine, K. A., . . . Balliet, D. (2013). Mate preferences do predict attraction and choices in the early stages of mate selection. *Journal of Personality and Social Psychology, 105*, 757–776.
- Liaw, A., & Wiener, M. (2015). randomForest: Breiman and Cutler's random forests for classification and regression (Version 4.6-12) [Computer software]. Retrieved from <https://cran.r-project.org/package=randomForest>
- Luo, S., & Zhang, G. (2009). What is attractive: Similarity, reciprocity, security or beauty? Evidence from a speed-dating study. *Journal of Personality, 77*, 933–964.
- McClure, M. J., Lydon, J. E., Bacus, J. R., & Baldwin, M. W. (2010). A signal detection analysis of chronic attachment anxiety at speed dating: Being unpopular is only the first part of the problem. *Personality and Social Psychology Bulletin, 36*, 1024–1036.
- McNulty, J. K. (2016). Should spouses be demanding less from marriage? A contextual perspective on the implications of interpersonal standards. *Personality and Social Psychology Bulletin, 42*, 444–457.
- Montoya, R. M. (2008). I'm hot, so I'd say you're not: The influence of objective physical attractiveness on mate selection. *Personality and Social Psychology Bulletin, 34*, 1315–1331.
- Murray, S. L., Holmes, J. G., & Collins, N. L. (2006). Optimizing assurance: The risk regulation system in relationships. *Psychological Bulletin, 132*, 641–666.
- Okbay, A., Beauchamp, J. P., Fontana, M. A., Lee, J. J., Pers, T. H., Rietveld, C. A., . . . Benjamin, D. J. (2016). Genome-wide association study identifies 74 loci associated with educational attainment. *Nature, 533*, 539–542.
- R Development Core Team. (2016). R: A language and environment for statistical computing (Version 3.3.2) [Computer software]. Retrieved from <https://www.r-project.org/index.html>

- Rusbult, C. E., & Van Lange, P. A. M. (2003). Interdependence, interaction, and relationships. *Annual Review of Psychology*, *54*, 351–375.
- Silver, N. (2012). *The signal and the noise: Why so many predictions fail – but some don't*. New York, NY: Penguin Press.
- Simpson, J. A., & Gangestad, S. W. (1991). Individual differences in sociosexuality: Evidence for convergent and discriminant validity. *Journal of Personality and Social Psychology*, *60*, 870–883.
- Spielmann, S. S., MacDonald, G., Maxwell, J. A., Joel, S., Peragine, D., Muise, A., & Impett, E. A. (2013). Settling for less out of fear of being single. *Journal of Personality and Social Psychology*, *105*, 1049–1073.
- Strobl, C., Malley, J., & Tutz, G. (2009). An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods*, *14*, 323–348.
- Tidwell, N. D., Eastwick, P. W., & Finkel, E. J. (2013). Perceived, not actual, similarity predicts initial attraction in a live romantic context: Evidence from the speed-dating paradigm. *Personal Relationships*, *20*, 199–215.
- Weigel, D., & Murray, C. (2000). The paradox of stability and change in relationships: What does chaos theory offer for the study of romantic relationships? *Journal of Social and Personal Relationships*, *17*, 425–449.