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

Samantha Joel[1]

Paul W. Eastwick[2]

Eli J. Finkel[3]

In Press, Psychological Science

[1] Department of Psychology, University of Utah

[2] Department of Psychology, University of California, Davis

[3] Department of Psychology and the Kellogg School of Management, Northwestern University

Samantha Joel
University of Utah
Department of Psychology
380 South 1530 East
Salt Lake City, UT 84112
Tel: (801) 213 3604
Email: samantha.joel@psych.utah.edu

Abstract

Close relationships theoretical perspectives and matchmaking companies 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 others (actor variance) and to be desired by others (partner variance), as well as desire for specific partners above and beyond actor and partner variance (relationship variance). In two speed-dating studies, romantically unattached individuals completed over one hundred traits and preferences identified by past research as relevant to mate selection. Participants then met one another in a series of four-minute speed-dates. Random forests models predicted 4-18% of actor variance and 7-27% of partner variance, but, crucially, they were unable to predict relationship variance using any combination of traits and preferences reported beforehand. 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*

Is Romantic Desire Predictable?

Machine Learning Applied to Initial Romantic Attraction

Achieving a high-quality romantic relationship is a goal with both evolutionary (Fletcher, Simpson, Campbell, & Overall, 2015) and practical consequence (Kiecolt-Glaser & Newton, 2001). Yet, the task of finding a suitable partner can be time-consuming and anxiety-provoking (Spielmann, MacDonald, Maxwell, Joel, Peragine 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, with companies striving 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 for the user to meet by feeding the questionnaire responses into an algorithm. Many companies claim to be able to match users with partners with whom they are especially likely to "click" upon first meeting (e.g., Chemistry.com, OKCupid.com). 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 theoretically far-fetched. Myriad perspectives in the close relationships and evolutionary psychological literatures suggest that outcomes such as relationship satisfaction and longevity follow from the conjunction of two partners' preferences, traits, and personal histories (e.g., Buss & Barnes, 1986; Byrne, 1961; McNulty, 2016; Campbell, Chin, & Stanton, 2016).

For romantic matching algorithms to be effective at all, one or more of the following three assumptions must be met: It must be possible to predict the emergence of romantic interest in the form of (a) who desires others on average (i.e., actor variance in desire), (b) who is desirable on average (partner variance), and (c) who uniquely desires whom (i.e., relationship variance; Eastwick & Hunt, 2014). If the first or second assumptions were true, an algorithm could help people form relationships by excluding exceptionally misanthropic (i.e., low actor effect) and/or undesirable (i.e., low partner effect) people from the group of eligible daters. But it is the third of these components—unique desire—that 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 our 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 stress-vulnerability-adaptation model (Karney & Bradbury, 1995) and the ReCAST model (Eastwick, 2016) highlight the chance, dyad-specific, 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, in press; Weigel & Murray, 2000). Consistent with these models, the strongest predictors of relationship outcomes (e.g., maintenance, dissolution) tend to be features of the relationship itself—like 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., if Laura and Ben share similar interests, they are no more or less likely to be attracted to each other; Luo & Zhang, 2009; Tidwell, Eastwick, & Finkel, 2013), and to idiosyncratic mate preferences (e.g., if Laura reports a preference for extraverted men and Ben reports that he is extraverted, Laura is no more or less likely to 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., Montoya, 2008; McClure, Lydon, Bacus, & Baldwin, 2010). 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 only be possible 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 machine learning method called random forests (Breiman, 2001; Liaw & Wiener, 2002). 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 non-linear 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 provides the most thorough and comprehensive test to date of the notion that romantic attraction can be predicted from self-reported traits and preferences.

Methods

In two samples of speed-daters, we used random forests (Liaw & Wiener, 2002) to predict romantic desire. As described briefly above, random forests are a machine learning technique 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 only accommodate 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, giving 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 non-parametric, meaning that 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, M_{age} = 19.6 years, SD_{age} = 1.0 years) who attended one of seven speed-dating events in 2005, and Sample B consisted of 187 undergraduate students (93 women and 94 men, M_{age} = 19.6 years, SD_{age} = 1.2

years) who attended one of eight 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 were recruited via on-campus flyers and emails to participate in a speed-dating study, with 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-minute online questionnaire that included a wide range of psychological constructs, including personality measures (e.g., the Big Five personality dimensions, attachment style, perceptions of 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), self-reported traits (e.g., warmth, physical attractiveness), and 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 the leading relationship scientist Caryl Rusbult in the late 1990s and early 2000s. Eli Finkel, a co-author on this report 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. When making decisions about which measures to include in the current study, we relied heavily on that Finkel study. Second, we reviewed the attraction literature from the 1960s-1980s and the evolutionary psychological literature on human mating from the 1990s-2000s and incorporated several individual-differences constructs from those literatures as well. The full 30-minute questionnaire was designed to be maximally comprehensive of these fields, and indeed, the constructs we prioritized are widely used (collectively cited 96,236 times as of March 1 2017; see Databases S1 and S2 for references) 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 paper, we included nearly all psychological constructs as predictors (182 constructs in Sample A; 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)?"). See Databases S1 and S2 on the Open Science Framework (OSF) for all measures collected at intake for Samples A and B, respectively: https://osf.io/4d3b9/. Items included as predictors in the present analyses are listed in Sheet 1 ("Items Selected") and items not included are listed in Sheet 2 ("Items Not Selected"), in the order in which they were collected from the participants.

Of the items included in present analyses, 8% of Sample A items and 19% of Sample B items have also been included in analyses reported in previously published papers (see Column M in Databases S1 and S2). Means, standard deviations, and ranges are also provided for each continuous measure. Variability was generally substantial across these measures: Across samples, most continuous variables have 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; cf. Li et al., 2013). Thus, there is little reason to expect that these variables would collectively fail to predict romantic desire a priori on the basis of their psychometric properties.

Approximately 1-2 weeks after completing the intake questionnaire, participants attended a speed-dating event in which they had a series of 4-minute speed-dates with ~12 opposite-sex individuals. Immediately following each speed-date, participants filled out a two-minute Interaction Record Questionnaire containing items that assessed their experiences on their most recent speed-date. In subsidiary analyses reported below, we used most of these constructs (18 in Sample A, 20 in Sample B) as predictors in the random forests models (see Databases S3 and S4 on OSF for all post-interaction measures collected for Samples A and B, respectively: https://osf.io/4d3b9/).

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 to 9 = Strongly Agree; Sample A α = .88, *M* = 5.04, *SD* = 2.11; Sample B α = .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 four-minute interaction—is comprised of 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 participant's actor effect and the partner's partner effect). 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 (SRM) analyses using the BLOCKO program (Kenny, 1998) that partitioned romantic desire into actor, partner, and relationship variance. These analyses revealed that a non-trivial percentage of romantic desire in the present samples can be attributed to each of these three sources (Table 1). Relationship variance was the largest source of variance, followed by partner variance, followed by actor variance; all three exceeded the "meaningful" threshold of 10% (Kenny, Kashy, & Cook, 2006). In other words, these reports are ideal for testing questions about the ability to predict actor, partner, and relationship variance, because all three are present in the dependent measure. (If anything, it might be easiest to predict relationship variance given that it is the largest source of variance.)

Table 1: BLOCKO variance partitioning

		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		

We next separated each romantic desire report (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/her speed-dating partners on average—by subtracting the romantic desire grand mean from the average of each participant's ~12 romantic desire reports. Second, we calculated partner desire—the extent to which the participant was liked by his/her speed-dating partners on average—by subtracting the romantic desire grand mean from the average of the ~12 romantic desire reports about that participant. Third, we calculated relationship desire—the extent to which the participant liked a particular partner above and beyond his/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 each romantic desire report. In the analyses below, we attempt to predict each of these three components separately.

Random Forests Analysis Strategy

For models predicting actor and partner desire, datasets 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, datasets were organized at Level 1, such that each observation was a dyad. Thus, each participant was represented ~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 analyses separately 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 1092 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 for R (Liaw & Wiener, 2002). For all analyses, we set "ntree" to 5000, meaning that each model was constructed from 5000 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; 2015). We constructed models using variable selection criteria at three levels of stringency. The “threshold” step of VSURF eliminates variables that fail to reduce the model’s error rate (liberal selection). The “interpretation” step of VSURF eliminates variables that fail 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 minimizes the number of predictors while maintaining predictive power (stringent selection). (For procedural details on how the VSURF R package selects predictors, see Genuer et al., 2010; 2015).

We also constructed models in which no selection criteria were used, such that all predictors were included in each model: see Supplementary Table 1. The amount of variance explained was substantially worse without the use of variable selection, suggesting that including many irrelevant predictors harmed the models’ predictive power. However, the amount of variance explained by the models varied little based on which selection criterion was used: see Table 2.

R syntax for all analyses is available on OSF: https://osf.io/4d3b9/.

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.

Table 2: Summary of primary random forests models predicting actor, partner, and relationship desire in Samples A and B

		Sample A			Sample B		
Dependent Measure	Variable Selection	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 (M)	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 (F)	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	*NA*	*NA*

Note. MSE: mean squared error; A lower MSE indicates that the model has a lower error rate. Actor desire: participant's responses. Partner desire: responses where the participant is the target. Relationship desire (M): a man's desire for a particular woman, beyond his actor effect and her partner effect. Relationship desire (F): a woman's desire for particular man, beyond her actor effect and his partner effect. Liberal variable selection eliminated only irrelevant variables, moderate variable selection kept moderately predictive variables, and stringent variable selection kept only the most predictive variables

Resulting models predicted approximately 5-18% of the variance in actor desire and 18-27% 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/responsiveness in a speed-date and one's own expected selectivity when choosing dates; see Supplementary Tables 2 and 3. In other words, people who see warmth as an attractive quality tended to experience greater attraction for their dates on average, and people seemed to have some insight into their own level of choosiness in the context of speed-dating. Consistent predictors of partner desire (i.e., the tendency to be desired by others) included participants' self-reports of their mate value and physical attractiveness: see Supplementary Tables 4 and 5. These results suggest that people have knowledge of their own attractiveness; people who self-reported high mate value and high physically 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 Supplementary Tables 6-9. However, predictors were not consistent across models; indeed, many of the relationship models explained a *negative* percentage of variance. The "percent variance explained" value is computed by the RandomForest package as

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

where MSE_{OOB} is the model's mean squared error and $\hat{\sigma}_y^2$ is the variance of the dependent observations (Liaw & Wiener, 2002). Therefore, a negative "percent variance explained" score 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 be desired (partner variance).

Training/Testing Analyses

An advantage of machine learning procedures such as random forests is that models that have been "trained" on one dataset can then be used to predict outcome measures in different dataset. Thus, these techniques are designed to answer questions about prediction in a truly a priori way. We next constructed additional models using data from only Sample A (the training data) and considering only the 87 predictors that were available in both datasets (shared variables are noted in the fourth column of Databases S1 and S2). 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, allowing us to generate predicted actor, partner, men's relationship, and women's relationship desire scores for Sample B. We then compared our generated desire scores to Sample B's actual desire scores to determine how well we were truly able to predict these dependent variables: see Table 3.

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	.88	15.48%	.88	.19**
Partner Desire	13	1.32	21.49%	1.26	.26***
Relationship Desire (M)	1	1.86	-2.19%	1.68	-.06, *ns*
Relationship Desire (F)	1	2.07	-0.56%	1.76	.02, *ns*

Notes. ** $p < .01$, *** $p < .001$

Selected predictors in each final, trained model are presented in Supplementary Table 10. The predicted actor desire scores for Sample B correlated with the *actual* actor desire scores for Sample B at $r = .19$, $CI_{95\%}[.05,.33]$, and the predicted partner desire scores correlated with the actual partner desire scores at $r = .26$, $CI_{95\%}[.12,.39]$. In contrast, men's predicted relationship desire scores for Sample B correlated with men's actual desire scores at $r = -.06$, $CI_{95\%}[-.12,-.002]$, and women's predicted desire scores for Sample B correlated with women's actual desire scores at $r = .02$ $CI_{95\%}[-.04,.08]$. 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 how much they will tend to desire others and how desirable they will be to others in a speed-dating context with some degree of accuracy. However, we cannot anticipate how much those individuals will uniquely desire each other in a speed-dating context with any meaningful level of accuracy.

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 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 post-interaction 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 post-interaction measures that participants completed following each speed date (e.g., perceived chemistry with the date, perceived intelligence of the date; see Database S3). For Sample B, the predictors were 20 post-interaction measures (see Database S4). In both samples, the only post-interaction 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 1092 rows. Analyses were conducted separately predicting male and female relationship desire, using the same analysis strategy used for the primary random forests models reported above.

Results are presented in Table 4. Unlike the original models constructed with background questionnaire measures, these models constructed with post-interaction measures predicted approximately 21-29% of male relationship desire and 16-24% of female relationship desire. The best predictor across both samples and both sexes was feelings of chemistry with the date (Supplementary Tables 11-14). 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.

Table 4: Random forests models predicting relationship desire in Samples A and B using post-interaction predictors

		Sample A			Sample B		
Dependent Measure	Variable Selection	Number of Predictors	MSE	Total Variance Explained	Number of Predictors	MSE	Total Variance Explained
Relationship Desire (M)	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 (F)	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%

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 41 predictors and 187 rows. Results are presented in Supplementary Table 15. People's post-interaction 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 participant's judgment of the dates' physical attractiveness and the participant's feelings of chemistry on their dates. The most consistent predictors of partner desire were the dates' judgment of the participant's physical attractiveness and the dates' feelings of chemistry with the participant.

As an alternative way to test the validity of the random forests method, we generated speed-dating datasets in which the romantic desire DV was generated from a combination of randomly generated actor effects (i.e., simulated preferences), randomly generated partner effects (i.e., simulated traits), and/or randomly generated relationship effects (i.e., simulated interactions of preferences and traits). In other words, in some of these datasets, romantic desire was a function of interactions between characteristics of the two partners, and in other datasets, romantic desire was only a function of characteristics of the actor and/or of the partner. Then, we used random forests to identify whether there are robust predictors of actor, partner, and relationship desire (just as in the primary random forests models above) from among the predictors that were actually used to create the desire DV (Supplementary Table 16). As expected, random forests were able to predict actor desire, partner desire, and relationship desire to the extent that the DV was originally comprised of 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 × trait interactions were pre-specified 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 (Supplementary Table 17). These results suggest that random forests are capable of detecting interaction effects where interaction effects exist. Thus, the inability of random forests to predict relationship variance in Table 2 is consistent with the possibility that relationship desire is not comprised of predictable interaction effects among background variables in our speed-dating data.

Discussion

The present research sought to predict initial romantic desire as accurately as possible across two speed-dating studies, using machine learning and over 100 self-report measures collected prior to the speed-dating events. We found that random forests could predict 4-18% of the variance in actor desire and 7-27% 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 Supplementary Tables 2 and 3), and be desired by others (e.g., sociosexuality, mate value; see Supplementary Tables 4 and 5). However, models drawing from these background measures were consistently unable to predict any variance in relationship desire: how much one person especially desired another person. Relationship desire was not predictable from background measures even though (a) it comprised the largest proportion of variance in the dependent measure (Table 1), and (b) random forests were successful at predicting relationship desire using relationship-specific measures collected post-interaction (16-29% of variance explained).

One possible interpretation of these findings is that background measures *could* be collected that would predict relationship-level desire, but relationship science has yet to reveal what they are. A related possibility is that relationship desire is comprised of a great many actor

× partner interaction effects—each 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 from 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 qualities that people initially find romantically desirable only weakly match the qualities that they articulate in the abstract (Eastwick et al., 2014).

The present findings only obliquely address 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 prior to 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., in press). 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 may exhibit matching by sociological factors such as age, SES, 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 likely prefer to date 20-year olds rather than 40-year olds, and vice versa). Further, the present study examined romantic desire experienced after a four-minute interaction. As people become better acquainted, it is unclear whether individual characteristics (assessed prior to an initial interaction) would become more predictive of relationship desire over time, or whether relationship desire remains predictable only from features of the relationship itself.

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 if we are to fully tackle this vexing and timeless question.

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 hypotheses presented in the current paper. 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 authors approved the final version of the manuscript for submission.

Acknowledgments

This research was supported by a postdoctoral fellowship awarded to the first author from the Social Sciences and Humanities Research Council of Canada. We thank Jacob Westfall and Tal Yarkoni for providing feedback during the development of this work.

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.*

Eastwick, P. W., Harden, K. P., Shukusky, J. A., Morgan, T. A., & Joel, S. (in press). Consistency and Inconsistency among romantic partners over time. *Journal of Personality and Social Psychology*.

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.

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., Simpson, J. A., Thomas, G., & Giles, L. (1999). Ideals in intimate relationships. *Journal of Personality and Social Psychology, 76,* 72-89.

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.

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. and Impett, E. A. (2012). Approach and avoidance motives and close relationships. *Social and 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. (2015). VSURF: An R package for variable selection using random forests. *The R Journal, 7,* 19-33.

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.

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 program]. Available: http://davidakenny.net/srm/srmp.htm

Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic data analysis.* New York: Guilford.

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.

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.

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News, 2,* 18-22.

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.

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.

Miller, R. (2015). *Intimate relationships.* Boston, MA: McGraw-Hill Education.

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. et al. (2016). Genome-wide association study identifies 74 loci associated with educational attainment. *Nature, 533,* 539-542.

Rusbult, C. E., & Van Lange, P. A. M. (2003). Interdependence, interaction, and relationships. *Annual Review of Psychology, 54,* 351-375.

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.

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.

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.

Silver, N. (2012). *The signal and the noise: Why so many predictions fail – but some don't.* New York: Penguin Press.

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.