Science Advances
AAAS

advances.sciencemag.org/cgi/content/full/4/8/eaap9815/DC1

Supplementary Materials for

Aspirational pursuit of mates in online dating markets

Elizabeth E. Bruch* and M. E. J. Newman

*Corresponding author. Email: ebruch@umich.edu

Published 8 August 2018, *Sci. Adv.* **4**, eaap9815 (2018)
DOI: 10.1126/sciadv.aap9815

This PDF file includes:

SUPPLEMENTARY INFORMATION

Section S1. Background literature

In the introduction, we presented two competing hypotheses of mate preferences: matching and competition. Here we expand on this discussion and provide an evaluation of previous efforts to study desirability in dating markets.

A variant of the matching hypothesis, originally proposed by Walster (*11*) and now considered a classic in the psychology literature, posits that all (or most) men and women do find the same attributes desirable and are attracted to the same potential partners, and they might as a result feel the urge to pursue someone "out of their league," as the competition hypothesis would suggest. They do not, however, either for fear of rejection, or simply to maximize their chances of success. Instead, they take into account their own desirability when deciding who to pursue (*11,36–40*) and only approach mates of desirability similar to their own (*34,35*). The net result, once again, is an assortative pairing of like with like.

Walster's work and subsequent studies building upon it assume that men and women have an overall desirability or "mate value," a score by which people, implicitly or explicitly, rank themselves and others. Our work has something in common with this approach in that, rather than relying on personal characteristics, we quantify desirability using empirical measures of who actually receives the most attention and from whom. This lends itself to the study of hierarchy in dating markets—our primary focus in this paper—by placing all actors on a single scale and allowing us to quantify concepts such as reach and desirability gap.

If one were to observe matching according to such overall desirability measures it would suggest either that people prefer partners of similar desirability or, as Walster hypothesizes, pursue partners of similar desirability for pragmatic reasons. In practice this is observed to some extent—most people pursue partners whose desirability is not too different from their own—but at the same time most people are "aiming high," with the average potential partner pursued having somewhat higher desirability than their pursuer. It is in this sense that we describe the observed behavior as a hybrid of matching and competition behaviors.

1.1 Previous work on desirability and matching

Previous studies have not provided a crisp formal definition of desirability or articulated a clear link between desirability and mate-seeking behavior. Related concepts have, however, arisen, particularly in the work of Fiore, Taylor, and coauthors (*12, 41*). It is useful to discuss their work in some detail, since they also make use of large-scale online dating data.

The most relevant study is that of Taylor *et al.* (*12*), in which the authors consider empirical evidence for Walster's matching hypothesis. They define a desirability score based on popularity, as measured by the number of first messages received on a dating site, and report a number of findings. First, they find that there is a weak but positive correlation between the popularity of message senders and receivers (R-values of 0.27 for messages sent by women and 0.37 for messages sent by men). They interpret this as evidence of popularity-based matching. Second, they assess whether people who match in this way are more successful, in terms of receiving replies, than those who do not. They compute the mean absolute difference between users' own popularity and the popularity of people they contacted. Their motivation for looking at absolute differences is that the matching hypothesis predicts that it should be equally disadvantageous to contact people who are either more or less desirable than yourself. They correlate each user's average success rate with their average absolute popularity gap, and find that this correlation is very close to zero (R-values of -0.01 for men and 0.00 for women). Taylor *et al.* interpret these results as indicating that both men and women tend to contact partners who are "in their league", but that their chances of getting a reply would not be affected if they did not.

This interpretation is in stark contrast to our own findings. We find that both men and women tend to message up the desirability hierarchy, and that there is a pronounced drop in

the probability of reply with increasingly positive desirability gap. There are a number of methodological differences between our approach and that of Taylor *et al.* that could account for this disagreement. To start with, Taylor *et al.* use simple popularity—number of messages received—as a proxy for desirability, which immediately introduces difficulties: if more desirable individuals receive more messages it implies that either (a) they are receiving a lot of messages from individuals less desirable than themselves, which would run counter to Taylor *et al.*'s claims that this is not happening, or (b) they are receiving messages from people of similar high desirability but such high-desirability people are sending more messages on average than low desirability people, which is the opposite of what we observe to be true.

If we nonetheless accept popularity as a measure of mate desirability, the results reported by Taylor *et al.* are not sufficient to prove the presence of matching in their data for two reasons. First, the existence of a correlation between the popularity of message senders and receivers, even a much stronger one than the authors find, cannot be used as evidence for matching, since correlation *R*-values of the type used by Taylor *et al.* are unaffected by uniform differences in popularity. In other words, one could achieve the same *R*-value if all individuals were messaging others more popular than themselves as one would if they were messaging others of the same popularity. The *R*-value is simply not sensitive to the absolute value of popularity and hence cannot be used as a measure of matching.[1]

Second, the authors' analysis of the relationship between desirability gap and reply probabilities conceals from view much of the dependence of replies on receivers' desirability. Recall that the authors measure the correlation between users' average absolute desirability gap and average rate of receiving replies. Use of the average gap obscures any variation of reply rate with desirability gap for an individual user, and, more importantly, the focus on *absolute* gap size means that the analysis cannot distinguish effects of messaging up versus down the desirability hierarchy. Our study reveals that people who message down the desirability hierarchy have a higher chance of reply than people who seek out partners of similar desirability, and people who

[1]Another study by the same team of authors (*41*) suffers from similar problems. They find that a multiplicative interaction between the popularity of message senders and receivers is slightly positively associated with the probability of reply. The multiplicative interaction suggests that reply rates are lowest when senders and receivers are both unpopular, and reply rates are highest when senders and receivers are both popular. Between those limits there are intermediate reply probabilities. This also does not demonstrate matching.

message up the hierarchy have a lower chance of reply than people who seek out partners of similar desirability. Taylor *et al.* by contrast conclude that there is no effect, but this appears to be an artifact of the way their analysis combines positive and negative gaps in a single measure, resulting in an average change in reply rate close to zero.

Our finding, based on the PageRank measure of desirability, is that individuals are in fact *not* matching on desirability. It is true that there is a positive correlation between desirability of sender and receiver, and most users message others of desirability not too dissimilar from their own, but there is also an offset, with most focusing on potential partners of higher rather than lower desirability.

In this sense our findings are different from those of Taylor *et al.*, and yet we still conclude that there is a hierarchy of desirability within the community we study and that the patterns of who pursues whom are strongly correlated with it. How can we make this statement in the absence of strong matching? The crucial observation is that most users message across a range of desirabilities, but receive replies only when they send messages to others who are of similar or lower desirability to themselves. Thus the observed behavior seems to be a hybrid of the traditional matching and competition models. On the one hand, people appear to be aware of their place in the desirability hierarchy and make their overtures accordingly, since on average they send messages to others who are not greatly more desirable than themselves. One could say therefore that a weak (and biased) form of matching is taking place.

This critique also highlights the danger of conflating matching with the existence of a hierarchy of desirability. While the matching hypothesis—as Walster formulated it—is related to hierarchy, the two do not necessarily imply the same behaviors. Most notably, the matching hypothesis implies that men and women do not engage in aspirational mate pursuit. While our analysis is largely descriptive in nature, our hope is that such rich description can improve social scientists' theoretical understanding of the behaviors that produce hierarchies in dating markets. Large-scale activity datasets such as those produced through online dating provide a unique opportunity to study human behavior at a high level of granularity. However, such data require our theories and models to move beyond the conceptual architecture developed for analyses of surveys and administrative data.

Table S1. User attributes for four metropolitan areas.

	New York		Boston		Chicago		Seattle	
	Men	Women	Men	Women	Men	Women	Men	Women
Total number of users	44009	50618	9113	9355	28635	23236	12721	9248
Ethnicity (%)								
Asian	8	11	4	6	3	4	7	9
Black	9	9	6	6	7	9	4	3
Hispanic	10	8	3	3	8	7	3	3
White	73	73	87	85	81	80	87	85
College degree (%)	92	96	70	80	63	71	64	68
Children at home (%)	5	6	7	10	7	10	15	17
Mean age	31.6	31.5	30.4	30.3	31.4	32	32.7	33.1
Mean messages sent	23.3	9.4	14.6	6.3	19	10.2	12.4	7.8
Replies received (%)	15	34	17	37	18	40	20	45

Section S2. Descriptive statistics

Table S1 gives a set of summary statistics for the male and female user populations in each of the four cities that are the focus of our study. The cities display a range of values of the ratio of men to women, New York having the largest fraction of women, followed by Boston, Chicago, and Seattle, in that order. Overall, the site has approximately 55 men for every 45 women. This slight excess of men is consistent with other studies of US online dating (*13, 15, 20*). In addition to their sex ratios, the cities differ in their overall market size and composition. New York City has the largest number of users, followed by Chicago, Seattle, and Boston. The average user is in their early 30s in all cities but there is modest variation in this and other demographic characteristics. Seattle users, for example, are slightly older and are more likely to have children living at home. Figure S1 shows the age distribution of men and women in each city. We see that New York has a surplus of women which is most pronounced among users in their mid twenties. The remaining cities all have a surplus of men, which is most pronounced in the later 20s and early 30s.

Table S1 also shows the average number of initial contacts made by men and women in each city and the percentage of those contacts that receive replies. As observed in other studies (*13, 15, 20*), men send significantly more messages than women. Overall they are

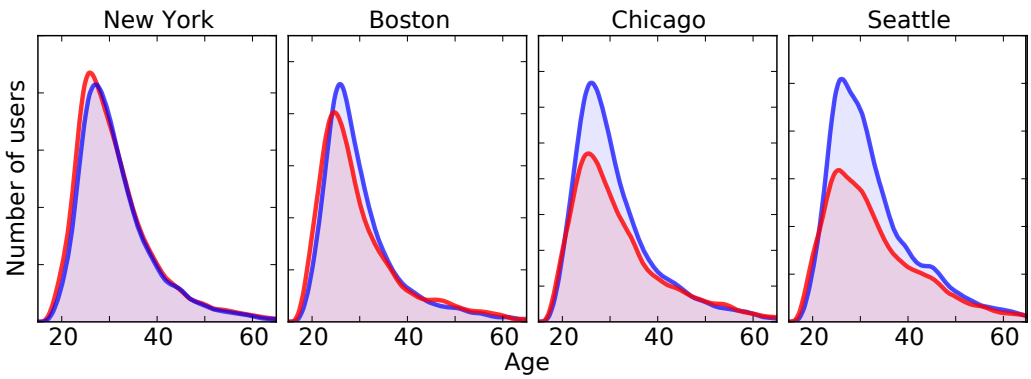

Fig. S1. Age distribution of men (blue) and women (red) in each city. Seattle and Chicago, and to a lesser extent Boston, have surpluses of men, the surplus being most pronounced for people around 28 years of age. New York city has a surplus of women, which is most pronounced among women in their mid-twenties. Note that because the total number of users varies across cities, the scale of the *y*-axis differs across the four panels.

responsible for 81% of initial contacts on the site, but men have a lower chance than women of receiving replies to their messages. This is not surprising: women may well reply less often precisely because they receive so many messages. The number of messages sent does, however, show some interesting variation between cities. Notice, for example, that among the cities studied men send the smallest number of messages and experience the largest reply rate in Seattle, which is unexpected since this is the poorest dating environment for men in terms of ratio of men to women. One might imagine that in cities where the sex ratio puts men at a disadvantage they would send more messages, in the hope of getting a reasonable number of replies. Moreover, the low number of messages sent by men in Seattle cannot be explained as a result of a larger fraction of inactive users, which might occur if male users become discouraged by the poor dating environment. As described above, only active users are included in the data, although it is possible that Seattle might contain a larger-than-usual number of users of low (but nonzero) activity level.

Section S3. Supplementary analyses

In this section we describe the statistical models used to create Figs. 2 and 4.

1.2 Desirability as a function of selected demographics

Figures for average desirability as a function of demographic characteristics shown in Fig. 2 of the main paper were calculated using fractional regression models of men and women's relative desirability as a function of their attributes. We specify robust standard errors to allow for clustering of observations within cities, while model interactions allow the effects of demographics on relative desirability to vary by city. The main effects refer to the values for Boston. Coefficient estimates from the fractional regression models are shown in table S2. In addition to those covariates included in the table, we control for other user attributes that might affect desirability: users' body type, whether they have children, and the type of relationship sought (e.g., short-term relationship, long-term relationship, or sex). A complete set of coefficients are available from the authors by request.

Table S2. Fractional regression of desirability on individual attributes—selected coefficients.

	Women			Men		
	Coef.	Std. error	*z*	Coef.	Std. error	*z*
Age	−0.055	0.014	−3.840	0.036	0.014	2.480
Chicago × Age	−0.019	0.016	−1.160	0.019	0.016	1.160
NYC × Age	−0.040	0.016	−2.530	0.030	0.016	1.860
Seattle × Age	−0.012	0.019	−0.660	−0.006	0.018	−0.310
Age^2	0.000	0.000	1.860	0.000	0.000	−1.990
Chicago × Age^2	0.000	0.000	0.940	0.000	0.000	−0.930
NYC × Age^2	0.000	0.000	2.260	0.000	0.000	−1.830
Seattle × Age^2	0.000	0.000	0.630	0.000	0.000	0.350
Black	−0.729	0.126	−5.770	0.227	0.142	1.600
Latino/a	−0.014	0.153	−0.090	0.278	0.173	1.600
White	−0.110	0.099	−1.120	0.492	0.114	4.330
Black × Chicago	0.301	0.146	2.070	0.057	0.162	0.350
Black × NYC	0.275	0.132	2.080	0.038	0.151	0.250
Black × Seattle	0.124	0.195	0.640	0.173	0.179	0.970
Latin × Chicago	0.015	0.171	0.090	0.000	0.190	0.000
Latin × NYC	0.004	0.159	0.030	−0.112	0.181	−0.620
Latin × Seattle	0.047	0.201	0.230	0.108	0.217	0.500
White × Chicago	0.182	0.116	1.570	0.045	0.132	0.340
White × NYC	0.014	0.103	0.130	0.101	0.120	0.840
White × Seattle	0.147	0.123	1.190	−0.019	0.133	−0.150
No college	−0.082	0.081	−1.010	−0.305	0.066	−4.590
Post-college	−0.132	0.045	−2.900	0.174	0.053	3.290
No college × Chicago	−0.050	0.090	−0.560	0.025	0.073	0.340
No college × NYC	−0.039	0.094	−0.420	0.005	0.077	0.060
No college × Seattle	−0.060	0.100	−0.600	0.009	0.080	0.110
Post-college × Chicago	0.090	0.053	1.710	0.006	0.061	0.110
Post-college × NYC	0.130	0.049	2.660	0.006	0.057	0.110
Post-college × Seattle	0.112	0.064	1.760	−0.122	0.069	−1.770
N		32832			31725	
Log-likelihood		−22247			−21662	

1.3 Message length and positivity by desirability gap

The top panels of Fig. 4 of the main paper show the relationship between desirability gap and message attributes. The predicted values of message length are derived from negative binomial regressions where the outcome is the total word count of the first message and the predictor variables are linear and quadratic terms for desirability gap. The predicted values of message positivity are derived from a fractional regression model where the outcome is the proportion of words in the message that are positively valenced (*22, 23*), and the predictor variables are linear and quadratic terms for desirability gap. Separate effects are estimated for each city via dummy variable interactions. The complete set of coefficients are shown in tables S3 and S4. The units of observation are first messages sent by a particular mate seeker to a potential match. Standard errors are robust to allow for clustering within mate seekers. To aid in ease of interpretation and presentation of results without excessive significant digits, the number of words in messages is divided by 100. (The values are scaled back up to their original levels in Fig. 4 of the paper.)

Table S3. Message length by desirability gap.

	Women			Men		
	Coef.	Std. error	*z*	Coef.	Std. error	*z*
constant	3.218	0.012	268.1	3.277	0.005	681.5
Chicago	−0.114	0.013	−8.600	−0.050	0.005	−9.560
NYC	−0.120	0.013	−9.520	−0.047	0.005	−9.280
Seattle	0.170	0.016	10.810	0.201	0.006	32.140
Desirability Gap	0.262	0.025	10.390	0.130	0.010	12.740
Desirability Gap × Chicago	0.113	0.028	4.010	−0.033	0.011	−2.930
Desirability Gap × NYC	0.118	0.027	4.390	−0.045	0.011	−4.140
Desirability Gap × Seattle	−0.040	0.033	−1.210	0.127	0.013	9.570
Desirability Gap^2	−0.134	0.050	−2.670	−0.031	0.020	−1.570
Desirability Gap^2 × Chicago	−0.133	0.056	−2.380	0.047	0.022	2.150
Desirability Gap^2 × NYC	−0.030	0.053	−0.570	−0.016	0.021	−0.730
Desirability Gap^2 × Seattle	0.049	0.066	0.750	−0.113	0.026	−4.350
N		188 774			1 285 568	
Log-likelihood		−784232			−5482604	

Table S4. Proportion of positive words in message by desirability gap.

	Women			Men		
	Coef.	Std. error	*z*	Coef.	Std. error	*z*
constant	−2.579	0.016	−165.5	−2.442	0.006	393.3
Chicago	0.123	0.017	7.030	0.053	0.007	7.890
NYC	0.043	0.016	2.590	0.014	0.007	2.190
Seattle	0.113	0.020	5.750	−0.014	0.008	−1.810
Desirability Gap	0.134	0.035	3.820	−0.102	0.013	−7.670
Desirability Gap × Chicago	−0.022	0.040	−0.550	−0.127	0.015	−8.780
Desirability Gap × NYC	0.023	0.037	0.620	−0.055	0.014	−3.940
Desirability Gap × Seattle	0.004	0.045	0.100	−0.011	0.017	−0.670
Desirability Gap^2	−0.001	0.067	−0.010	−0.133	0.027	−5.010
Desirability Gap^2 × Chicago	0.072	0.076	0.940	0.197	0.029	6.850
Desirability Gap^2 × NYC	−0.024	0.071	−0.330	0.259	0.028	9.290
Desirability Gap^2 × Seattle	0.038	0.086	0.440	0.133	0.033	4.010
N		188 774			1 285 568	
Log Likelihood		−50732			−367246	

1.4 Reply probabilities by message length and positivity

The bottom panels of Fig. 4 of the main paper show expected reply rates as a function of message length and positivity. These values are derived from logistic regressions describing how the probability of receiving a reply to an initial contact varies with message length and the percent of words in the message that have positive connotations (*22*, *23*). Because online dating site users, especially women, tend to write longer and more positive messages to more desirable partners, we control for the desirability gap between sender and receiver. We also allow for potentially nonlinear effects of desirability, message length, and message positivity on the probability of receiving a reply. Separate effects are estimated for each city via dummy variable interactions. The complete set of coefficients are shown in tables S5 and S6. The units of observation are first messages sent by a particular mate seeker to a potential match. Standard errors are robust to allow for clustering within mate seekers. To aid in ease of interpretation and presentation of results without excessive significant digits, the number of words in messages is divided by 100. (The values are scaled back up to their original levels in the article figures.)

Table S5. Probability of reply by message length, conditional on desirability gap.

	Women			Men		
	Coef.	Std. error	*z*	Coef.	Std. error	*z*
constant	−0.130	0.031	−4.170	−1.276	0.014	−91.220
Chicago	−0.077	0.034	−2.270	−0.053	0.015	−3.490
NYC	−0.321	0.033	−9.770	−0.218	0.015	−14.770
Seattle	0.234	0.040	5.880	0.046	0.018	2.570
Desirability gap	−0.735	0.051	−14.410	−0.500	0.023	−21.390
Desirability gap × Chicago	0.104	0.057	1.830	0.116	0.026	4.530
Desirability gap × NYC	0.019	0.054	0.350	0.025	0.025	1.020
Desirability gap × Seattle	−0.030	0.068	−0.450	0.210	0.030	6.940
Desirability gap^2	0.226	0.101	2.240	0.285	0.045	6.260
Desirability gap^2 × Chicago	0.213	0.112	1.900	−0.283	0.050	−5.660
Desirability gap^2 × NYC	0.280	0.107	2.620	−0.293	0.048	−6.060
Desirability gap^2 × Seattle	0.100	0.134	0.750	−0.212	0.058	−3.640
Number of words	−0.140	0.111	−1.260	−0.059	0.041	−1.440
Number of words × Chicago	0.279	0.118	2.370	−0.040	0.044	−0.900
Number of words × NYC	0.111	0.117	0.950	0.054	0.042	1.270
Number of words × Seattle	0.271	0.125	2.170	0.243	0.047	5.210
Number of words^2	0.087	0.058	1.500	0.035	0.015	2.310
Number of words^2 × Chicago	−0.086	0.059	−1.450	0.001	0.016	0.060
Number of words^2 × NYC	−0.002	0.061	−0.040	−0.036	0.015	−2.350
Number of words^2 × Seattle	−0.078	0.060	−1.300	−0.039	0.016	−2.460
N		188 774			1 285 568	
Log-likelihood		−126 679			−637 918	

Table S6. Probability of reply by percent of positive words, conditional on desirability gap.

	Women			Men		
	Coef.	Std. error	*z*	Coef.	Std. error	*z*
constant	−0.167	0.030	−5.550			
Chicago	0.009	0.033	0.290	−0.057	0.016	−3.610
NYC	−0.288	0.032	−9.110	−0.223	0.015	−14.660
Seattle	0.374	0.040	9.320	0.125	0.019	6.610
Desirability gap	−0.742	0.051	−14.580	−0.503	0.023	−21.540
Desirability gap × Chicago	0.124	0.057	2.180	0.111	0.026	4.320
Desirability gap × NYC	0.029	0.054	0.540	0.025	0.025	1.000
Desirability gap × Seattle	−0.006	0.068	−0.090	0.223	0.030	7.380
Desirability gap^2	0.227	0.101	2.250	0.279	0.045	6.150
Desirability gap^2 × Chicago	0.207	0.112	1.850	−0.276	0.050	−5.520
Desirability gap^2 × NYC	0.278	0.107	2.600	−0.285	0.048	−5.890
Desirability gap^2 × Seattle	0.096	0.133	0.720	−0.213	0.058	−3.660
% Positive words	0.003	0.004	0.850	−0.006	0.002	−3.830
% Positive words × Chicago	−0.005	0.004	−1.320	−0.001	0.002	−0.870
% Positive words × NYC	−0.003	0.004	−0.910	0.002	0.002	1.220
% Positive words × Seattle	−0.013	0.005	−2.700	−0.001	0.002	−0.230
% Positive words^2	0.000	0.000	−0.610	0.000	0.000	1.690
% Positive words^2 × Chicago	0.000	0.000	0.430	0.000	0.000	1.540
% Positive words^2 × NYC	0.000	0.000	1.480	0.000	0.000	−0.670
% Positive words^2 × Seattle	0.000	0.000	1.510	0.000	0.000	−1.070
N		188 774			1 285 568	
Log Likelihood		−126 705			−637 843	