COMPATIBILITY OF MATING PREFERENCES

Halk O. Bingol and Omer Basar
Bogazici University, Istanbul, Turkey

Mating is important for evolution. Since mating requires agreement of both parties, although men and women have different preferences in partner selection, there should be compatibility in these preferences. How compatible are the preferences of male and female in a given property such as age, height, education and income?

We have investigated the mating preferences of females and males on a large online dating site. We use the most restricted definition of mating in our data which reduces the population size to $N = 44,253$. We confirm that females and males have different mating preferences. Females prefer taller and older males with better education and higher income. Males prefer just the opposite. Our findings indicate that these differences complement each other. A simple evolutionary model is developed which produces similar results.

Keywords: mating, mate selection, mating preferences, parental investment, gender compatibility, evolution, online dating

INTRODUCTION

There has been a long debate on how different male and female are [1–12]. Gender difference hypothesis claims that males and females are very different in their personalities, abilities, interests, attitudes and behavioral tendencies [2–10]. Some recent findings support this hypothesis. Men and women are different in many ways including sexual contacts [5], in brain imaging studies [7], performance in Mathematics (possibly due to sociocultural reasons) [8], or chess [10], even online games [9]. On the other hand, some investigations support the gender similarity hypothesis which claims that the difference is not so big as one expects [11, 12].

One possible evolutionary explanation of the gender difference is the hunter-gatherer theory of spatial sex differences proposed by Silverman and Earl [13]. It claims that there has been division of labor between men and women as early as the time of hunter-gatherers. Males are primarily hunters and females primarily foraged. This affects the cognitive development since "tracking and killing animals entail different kinds of spatial problems than does foraging for edible plants; thus, adaptation would have favored diverse spatial skills between sexes throughout much of their evolutionary history". Hence, it calls for different spatial skills such as mental rotations, map reading, maze learning for male, and ability to learn, recognize, remember spatial configurations of objects, and their spatial relationships for female.

MATING AND PARENTAL INVESTMENT

Mating is important for evolution. In many species, it has been observed that male and female have different strategies in mate selection.

We need an evolutionary theory to explain differences in mating strategies. One such theory is Trivers’ parental investment theory which is based on parental investment [14]. He carefully defines parental investment as “any investment by the parent in an individual offspring that increases the offspring’s chance of surviving (and hence reproductive success) at the cost of the parent’s ability to invest in other offspring”. Therefore, evolution calls for parental support, since offsprings, who have support from parents, have better chance to reproduce.

Parental investment is quite uneven between male and female in many species [14]. Therefore both genders evolutionarily developed mating strategies which are clearly different [3, 15–17]. Human female makes mandatory high investment in offspring compared to male, if one considers nine months of gestation, childbirth, lactation, nurturing. Therefore, she looks for supporting male in her mate selection. She prefers a male who not only have the resources to support her but also willing to commit these resources to her. That explains female preference for long-term commitment. On the other hand, human male can choose quantity. He is reluctant to long-term commitment, partly due to parental uncertainty, that is, he is not hundred percent sure that the child carries his genes. He has a tendency for short-term relations, which increase his chances to reproduce offsprings. This quality versus quantity paradigm is clearly a conflict that has to be resolved. Female, who invest more in offspring, should be more choosy selecting mate (intersexual attraction) and male, who invest less, should compete to access the opposite sex (intrasexual competition) [14].

PROPERTIES IN MATE SELECTION

Properties, that increase the chance of mating, become crucial in this respect [3, 13–21]. In terms of evolution, (i) fertility, i.e., immediate probability of conception, and (ii) reproductive value i.e., future reproductive potential, are the top two properties for both gender [17]. They are age related. Clearly, health is also very important. So young healthy mate should be the choice in all species.
Human male just does that. But human female has other issues therefore being young is not enough. She is looking for male that will provide parental support for her offspring, that is, he should have the resources and (ii) willingness to commit these resources to her offspring. In order to collect resources he needs time. Therefore he may not be that young after all. While female is busy providing parental investment to her offspring, she expects male to provide food and shelter. He is also expected to protect her offspring and her. Hence physically masculine male should be preferred. So we expect that younger female and “superior” male partners.

Empirical evidences support these deductions [3, 6, 13, 15, 18–21]. Investigations on partners has revealed that male is older [18], taller [19–21] than female. This is a universal pattern across cultures [6, 15].

Since mating requires agreement of both parties, although men and women have different preferences in mate selection, there should be compatibility in these preferences. We can ask the following question. How compatible are the preferences of two genders in a given property?

We will use the data obtained from an online dating site. First we carefully define mating in our data set. Then we aggregate the properties of partners that an individual selects as mate. Finally we search for patterns in the properties for mating behavior.

**METHOD**

**Data Set**

We investigate the data of a large online dating site for compatibility of mating preferences [22]. There are 4,500,000 registered users in total. More than 3,000 new users register daily. A user stay in the system for 3 months, on the average. Many of them come back, later; sometimes as a new user. The daily activity is also quite large such as 50,000 user logins, 500,000 massages, and 20,000 votes.

**Definition of Mating**

A typical online dating system enables its user to find a partner that best matches one’s desires. Each user defines his user profile. An initiator, predominantly male, selects a potential partner by examining her profile and sends her a message. If there is a positive response from the receiver, then more messages are exchanged which hopefully leads to a face-to-face meeting.

When do we say that male and female are mating? Online dating site has lots of information about the virtual world, but there is usually no information whether the male and the female really make a mating partner in the physical world. Any action in an online dating site is clearly an attempt for mating but is it sufficient to be considered as mating? For example, just sending a message, getting a message in response, or some more exchange of messages should not be enough in the world of online dating which is full of these.

Therefore we select the most restricted criteria of mutual interest, that is available in our data set, which is virtual gifts [22]. Receiving a virtual gift, which is usually a picture of a flower, is considered a “value” in this virtual society. We have even observed that some user sent virtual gifts to themselves. This value is probably due to (i) the virtual gifts one receives is visible to all, (ii) they are not free, i.e., one has to purchase virtual gifts in order to sent, and (iii) only qualified users can sent virtual gifts. Since unpaid male members are not qualified to sent gifts, able to sent gifts may be considered as an indication of wealth.

There are 276, 210 male and 483, 963 female users that are qualified to send virtual gift in the system. Among those, only 29, 274 male and 14, 981 female, in total $N = 44, 253$, users reciprocally exchange virtual gifts. Hence we define a pair as (mating) partners if they not only exchange messages, but also send at least one gift to, and receive at least one gift from each other.

Note that this definition is based on actual behaviors of users in the online dating site. We have the “actual” partners, that is, they mutually agree to “mate” as far as we can trace in our online dating site, rather than a “theoretical” partner one wishes to have as he answers the questions of a survey as in the case of ref [15]. We have such theoretical data in user profiles, too. Users specify what properties, such as the age, height, they look for in their potential partners. This data is noisy. Users are not consistent. They claim something and does something else. For example, someone claims that he prefers women taller than 170 cm but does not hesitate to be partner with a 160 cm. Such behavior is clearly difficult to register in questionaries. In this respect our actual data deserves special attention.

**Properties of the Mate**

Once we have identified the partners, we investigate the properties of the mate. As expected, user $i$ becomes partner with many others as time goes. Each partner of $i$ may have different value for property $p$. The average of the properties of the partners of $i$ is given as

$$\bar{p}_i = \frac{1}{|C_i|} \sum_{j \in C_i} p_j$$

where $C_i$ is the set of users that $i$ partnered with, and $p_j$ denotes the property $p$ as it is defined in user $j$’s profile. We interpret this as user $i$ has a tendency to select
partners having value of $p_i$ in property $p$. Hence, we call $p_i$ as the preferred value for $i$. Instead of using the preferred value directly, we compare one’s own value to the preferred value that one looks in his partners. The preferred difference of $i$, in property $p$, is defined as

$$\Delta p_i = \overline{p_i} - p_i.$$ 

Note that $\Delta p_i$ can be positive or negative. If $\Delta p_i$ is around 0 then the user prefers partners with similar properties with him, i.e. homophily [23, 24]. For example, if user $i$ has a tendency for taller partner in her selection, then $\Delta p_i$ in height would be positive.

**Preferred Difference Distributions**

We can extend these concepts from individual $i$ to a group of people. Then, frequency of people with the same preferred difference makes a probability distribution, which we call preferred difference distribution. Having all women as one group, and all men as another group, we obtain two preferred difference distributions $f(x)$ and $m(x)$ of females and males, respectively.

**RESULTS**

**TABLE I: Comparison of male and female distributions**

| Property     | Averages | Standard Deviations | Compatibility |
|--------------|----------|---------------------|---------------|
|              | $\mu_m$  | $\mu_f$  | $\sigma_m$ | $\sigma_f$ | $\rho$ |
| Height (cm)  | -11.12   | 11.37   | 6.76     | 7.09      | 0.90  |
| Education (bin) | -0.36      | 0.34     | 1.35    | 1.40      | 0.92  |
| Age (year)   | -2.90    | 2.74     | 5.06    | 5.23      | 0.94  |
| Income (bin) | -0.93    | 0.99     | 1.28    | 1.32      | 0.95  |

The statistical parameters of the preferred differences in age, height, education and income are given in Table I. Columns $\mu_m$, $\mu_f$, and $\sigma_m$, $\sigma_f$ are the averages and standard deviations of males and females, respectively. The distributions of the preferred differences are given in Fig. 1. We first focus on the averages, and leave the discussion of distributions and their compatibility later.

**Average of Preferred Differences**

In all four properties in Table I, there is a distinct pattern. The averaged preferred differences for males, $\mu_m$, are all negative and that of females are all positive. This observation indicates that in all four properties, whatever the metric is used to measure the property, males prefer “inferior” females and females prefer “superior” males compared to themselves.

**Age.** According to evolutionary theories, we expect to see younger female and older male in partners. Our findings confirms that. We observe that, on the average, males mate with females 2.90 years younger than themselves, and females mate with males 2.74 years older. Our findings are in agreement with ref [15], which reports that males prefer 2.66 years younger, females prefer 3.42 years older mate than themselves.

**Height.** Our findings on preferred difference in height, given in Fig. 1(a), agree with the previous work. People usually interact with people who have similar characteristics [23, 24]. For example no drastic height differences between partners are observed. That is, tall male partners with tall female, and short with short [19, 21]. Although there are males prefer females 30 cm shorter or 10 cm taller in Fig. 1(a), they are rare. Majority are accumulated around the average which is a manifestation of homophily in height.

Yet, there are distinct differences between the preferences of male and female when it comes to partner selection, such as height [19–21]. Male is usually taller then female in partners, which is called male-taller norm [19]. The averages in Table I agree with the male-taller norm. On the average, male prefers female 11.12 cm shorter. Similarly, on the average females prefer male 11.37 cm taller.

**Income and Education.** We observe similar pattern in income and education, too, namely, males prefer negative and females prefer positive differences. Here the numbers cannot be compared with other works directly since users are asked to select one bin out of many bins which are organized in a consistent but an arbitrary way [22]. They are consistent in the sense that the larger the bin number, the more educated or higher income. The bins in education are related to the number of schooling years such as graduate of primary school, or of college. The bins in income field represent monthly income such as bin 2: 500 $< x < 1000, \text{ bin 3: } 1000 < x < 2000$.

**Distribution of Preferred Differences**

Note that the absolute values of average preferred difference of males and females are very close to each other in Table I. The standard deviations are also very close to each other. These cannot be a coincidence and deserves further study.

Fortunately, we have not only the averages and standard deviations but also the distributions, that we can further investigate. Distributions of preferred differences in height, age, education, and salary are given in Fig. 1. The bell-shaped curves of male and female distributions are resemble to each other. One notices that female curves are right-shifted while male curves are left-shifted.
FIG. 1: (Color online) Preferred difference distributions in height, age, education, and income.

In order to get better understanding, consider a simplified example given in Fig. 2. Note that females that prefer $\Delta p = x$ matches with males that prefers $\Delta p = -x$. Therefore, we should not directly compare the distribution $f(x)$ of females with $m(x)$ of males. We should compare $f(x)$ with $m(-x)$, that is, the symmetric graph with respect to the $y$-axis. We make a reasonable assumption that there are equal number of men and women. Then $\min\{f(x), m(-x)\}$ of the women who prefer $\Delta p = x$ are matched. Then, the compatibility of two distributions can be measured by means of the ratio of matched women given as

$$\rho = \sum_x \min\{f(x), m(-x)\}$$

where summation is taken over all possible values of $x$. This is a well-defined metric since the ratio of matched women is equal to that of men.

In Table I, the properties are listed in ascending order in compatibility. Height is the property with the lowest compatibility. Even for this case, 90% of the population can find a satisfying partner. Interestingly, income has the highest compatibility and age comes next.

MODEL

A simple evolutionary model can explain high compatibility values. As an abstraction, we consider agents with one property only, such as age.

Agents. Agent $i$ represented by a 4-tuple $(g, p_i, p_i^{\min}, p_i^{\max})$, called genotype, where gender $g$ is binary number with 0 being female, and $p_i, p_i^{\min}, p_i^{\max} \in [a, b]$. $p_i$ is the own property and $[p_i^{\min}, p_i^{\max}]$ is the mating range of the agent. Agent $i$ accepts to mate with opposite gender agent $j$ if $j$ is in the mating range of $i$, that is, $p_j^{\min} \leq p_j \leq p_j^{\max}$. For example female $(0, 2, 4, 5)$ accepts male $(1, 4, 3, 4)$ but the male does not accept the female.
Generations. We start with a generation of $N$ female and $N$ male agents. We let agents get chance of reproduction by randomly picking a female and a male to meet. Some of these meeting produces a child. After $M$ meetings, the parents are removed and the children become the next generation. Note that the number of girls or boys is usually different than $N$. If the population of any gender exceeds $N$, we randomly eliminate some so that every generation has no more than $N$ females and $N$ males. If the population of a gender is less than $N$, we do not do anything to increase it to $N$. Of course, if the population of one gender becomes zero, then the system stops reproducing.

We have to initialize the very first generation. The rest of the generations are driven by the system. We initialize female and male agents differently. In order to initialize an agent, we draw three numbers from a uniform random distribution of $[a, b]$ and order them so that we have $p_1 \leq p_2 \leq p_3$. Then if the agent is female, set its genotype to $(0, p_1, p_2, p_3)$, otherwise to $(1, p_3, p_1, p_2)$.

Note that females accepts males that have higher value than her property since we use the minimum of the numbers for her own property. Similarly, males accepts females with lower property. So our model starts with agents that are agree with “female prefers superior, male prefers inferior” paradigm. Since child gets exactly the same genotype of either mother or father, in all generations this paradigm is preserved.

Initialization. We obtain the first generation by repeating agent initiation process $N$ times for female and $N$ times for male. Hence we expect that the initial generation has quite a variety of genotypes, possibly $2N$ different genotypes. Since there is neither mutation nor recombination of genotypes of mother and father, there is no way that the genotype variety of the system can increase. On the contrary, it can reduce when the last member of a genotype is failed to reproduce.

Observations. Simulations of the model reveal that genetic variation is reduced from generation to generation. Agents that cannot find mate are eliminated from the system. Agents with genotype that fits the population survives. The compatibility of the population increases from generation to generation. A sample runs of the model is given in Fig. 3. It is observed that the system converges to the compatibility of 1 most of the time. Convergence is quite fast. For $N = 100$, 200 generations are usually sufficient. Of course, there are some realizations that extinct but they are rare. Interestingly, in many realizations genotype pool reduces to one female and one male genotype.

DISCUSSION

In real life, men and women behave differently. Our findings show that the virtual world of online dating is
another manifestation of this difference. (i) While male prefers women with lower qualifications in every property that we have investigated, women just do the opposite. (ii) Interestingly, the preferences of men and women match to each other so that the number of dissatisfied is minimized. Due to lack of space, we do not report here but we have also observed similar findings in body mass index and body type, too [22].

One needs to be careful on a number of issues in a study like that. (i) One has to keep in mind that the findings could be culture dependent. (ii) The profile is based on user’s claim, that is, it may be misleading. On the other hand, stretching the properties too far would not be a good strategy since unfaithful declaration, such as declared as slim while being obese, would be an obstacle to further the relationship when the time comes to meet face-to-face [25, 26]. So we assume that users are closed to what they claim to be. (iii) Privacy is the most important issue for such an investigation. In this study no data left the company. All the data processing is done at their site. No individual personal information is used. Only statistical data such as given in Fig. 1 is shared with us.

This work was partially supported by Bogazici University Research Fund, BAP-2008-08A105, by the Turkish State Planning Organization (DPT) TAM Project, 2007K120610, and by COST actions MP0801 and TD1210. The authors would like to thank Adil Saribay for his assistance with psychology literature.

[1] C. Darwin, The descent of man, and selection in relation to sex (John Murray, London, 1871).
[2] D. Tannen, You just don’t understand: Women and men in conversation (Virago London, 1991).
[3] D. M. Buss, The evolution of desire: Strategies of human mating (Basic books, 2003).
[4] J. Gray, Men are from Mars, women are from Venus: Practical guide for improving communication (Zonder- van, 2009).
[5] F. Liljeros, C. R. Edling, L. a. Amaral, H. E. Stanley, and Y. Aberg, Nature 411, 907 (2001).
[6] D. P. Schmitt, A. Realo, M. Voracek, and J. Allik, J. Pers. Soc. Psychol. 94, 168 (2008).
[7] C. J. Cela-Conde, F. J. Ayala, E. Munar, F. Maestú, M. Nadal, M. a. Capó, D. del Río, J. J. López-Ibor, T. Ortiz, C. Mirasso, and G. Marty, PNAS 106, 3847 (2009).
[8] J. S. Hyde and J. E. Mertz, PNAS 106, 8801 (2009).
[9] M. Szell and S. Thurner, Scientific Reports 3, 1214 (2013).
[10] A. Blanch, A. Aluja, and M.-P. Cornadó, Pers. Individ. Dif. 86, 117 (2015).
[11] J. S. Hyde, Am. Psychol. 60, 581 (2005).
[12] E. Zell, S. R. Teeter, and E. Zell, Am. Psychol. 70, 10 (2015).
[13] I. Silverman and M. Eals, in The Adapted Mind Evolutionary Psychology and the Generation of Culture, edited by J. H. Barkow, L. Cosmides, and J. Tooby (Oxford University Press, 1992) pp. 533–549.
[14] R. L. Trivers, in Sex. Sel. Descent Man 1871–1971 (Aldine Publishing Company, 1972) pp. 136–179.
[15] D. M. Buss, Behav. Brain Sci. 12, 1 (1989), arXiv:arXiv:1008.2616v2.
[16] D. M. Buss and D. P. Schmitt, Psychol. Rev. 100, 204 (1993).
[17] D. M. Buss, Psychol. Top. 15, 239 (2006).
[18] D. T. Kenrick and R. C. Keefe, Behavioral and Brain Sciences 15, 75 (1992).
[19] J. S. Gillis and W. E. Avis, Personality and Social Psychology Bulletin 6, 396 (1980).
[20] B. Pawlowski, R. Dunbar, and A. Lipowicz, Nature 403, 156 (2000).
[21] V. Swami, A. Furnham, N. Balakumar, C. Williams, K. Canaway, and D. Stanistreet, Personality and Individual Differences 45, 395 (2008).
[22] H. O. Bingol and O. Basar, arXiv preprint (2012), arXiv:arXiv:1211.1035v1.
[23] M. McPherson, L. Smith-Lovin, and J. M. Cook, Annual Review of Sociology 27, 415 (2001).
[24] D. Centola, Science 334, 1269 (2011).
[25] G. Norcie, E. De Cristofaro, and V. Bellotti, LNCS 7862, 149 (2013), arXiv:arXiv:1303.4155v2.
[26] N. B. Ellison and J. T. Hancock, IEEE Secur. Priv. 11, 84 (2013).