Do People Actually Know What They Want? Predicting Romantic Matches from Preferences & Behavior

By Cheri Chu
(shared dataset with Cecilia Kuang and Rhea Remesh)

It is common to hear people discussing theories about what they value in a romantic partner, often listing a multitude of attributes such as attractiveness, ambition, or shared values. However, do these stated preferences actually reflect the cognitive processes that drive real-world romantic decision-making? Do individuals truly understand the factors that might influence their own choices, and can these self-reported preferences predict actual behavior in social interactions? To investigate, we analyzed data from a speed-dating study in which participants reported their demographic characteristics, the relative importance they assigned to various partner attributes, ratings of their partner after the dates, and match outcomes.

1. We ask whether it is possible to predict a mutual match using logistic regression models based on (1) demographic information, and (2) the participant and partner stated attribute weights before the dates.
    a. Additionally, which features/attributes are most correlated with a match if given more weight?
2. We also apply decision tree models to both sets of predictors: do people's stated theories about what they value align with what actually predicts their behavior after interacting?

Together, we hope to assess the structure of the decision-making process of romantic choice, as well as predictive accuracy. In understanding this, we may be able to suggest implications for theories of self-knowledge, decision-making, or social cognition.

Dataset Description

The dataset consists of observations from a structured speed-dating event, where 551 unique participants rotate through a series of short, face-to-face interactions. Each row in the dataset represents one interaction between two individuals (one "date"). For every pair, both participants and partners provided their personal theories on what is most important when looking for a partner before the interaction. After the interaction, both participants and partners rated each other on a scorecard. The dataset also includes information about their demographic profile, as well as whether they "matched" with their partner after the date.

The data include around 195 variable columns, so we decided to only focus on a select handful:

Demographic variables:	
Participant	**Partner**

Gender (only shown once per row, 0/1)	–
age	age_o
race	race_o

‘Dating theory’ variables: Out of a point system where total points equal 100, both participants and partners weigh what they think is important to them when looking for a partner:	
Participant weights out of 100:	**Partner** weights out of 100:
attr1_1 (The importance of attractiveness)	pf_o_att
sinc1_1(The importance of sincerity)	pf_o_sin
intel1_1(The importance of intelligence)	pf_o_int
fun1_1(The importance of fun)	pf_o_fun
amb1_1(The importance of ambition)	pf_o_amb
shar1_1(The importance of shared interests)	pf_o_sha

Scorecard variables: The participants and partners rate each other out of 10 after the 4-minute date:	
Participant ratings out of 10:	**Partner** rankings after that SAME date.
attr (attractiveness)	attr_o
sinc (how sincere)	sinc_o
intel (how intelligent)	intel_o
fun(how fun)	fun_o
amb(how ambitious)	amb_o
shar(how similar their interests were)	shar_o

The key outcome variable is match: coded as 1 if *both* participants said “yes” to seeing each other again, and 0 otherwise.

Preprocessing

The researchers collected the data in waves, or groups of participants. Due to a change in the experimental process across waves, we decided to use only the first 1-5 waves of the experiment that used the 100-point weight system. We believed this specific system provides a clearer representation of participants' dating theories because it forces them to distribute weight across all attributes, rather than assigning maximum values to all criteria. As a result, participants must make trade-offs and reflect carefully on their answers.

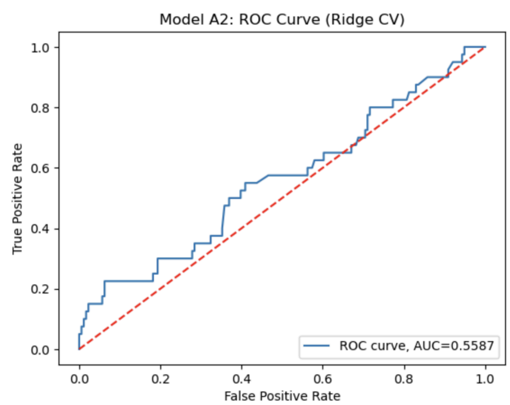
For exploratory data analysis (EDA), we selected only the variables mentioned above for the first 1-5 waves of the experiment. For the models, we divided the data into three datasets: Dataset A(demographic variables), Dataset B (dating theory variables), and Dataset C(scorecard data).

For all model datasets, we restricted to the female participants only. Each row represents one interaction, and both the participant's and partner's answers are recorded in that single row. Since this dataset was only collected on same-sex pairings, this would mean that attr1_1 would correspond with the woman's responses, and pf_o_att would correspond with the man's responses. Finally, it is worth noting that many participants were college-aged, and a large population of participants identified as White or Asian. Since we did not control for these imbalances, they may limit the generalizability of the results.

Results

Model A(Demographics Only):

- Ridge Logistic Regression: Accuracy : 0.819, AUC = 0.500

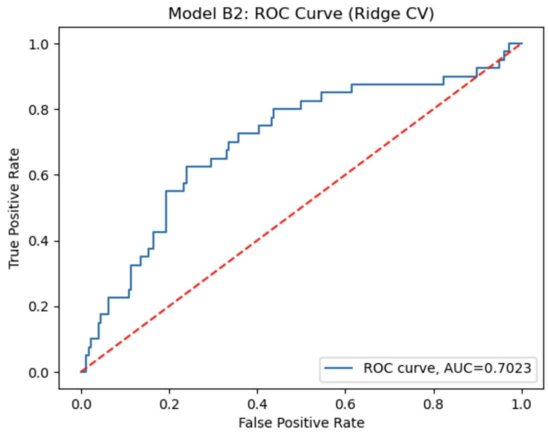


- Lasso Logistic Regression: Accuracy : 0.815, AUC = 0.500

Both models performed similarly, Lasso shrunk all coefficients to 0.

Model B(Preference Weights):

- Ridge Logistic Regression: Accuracy: 0.653, AUC = 0.734

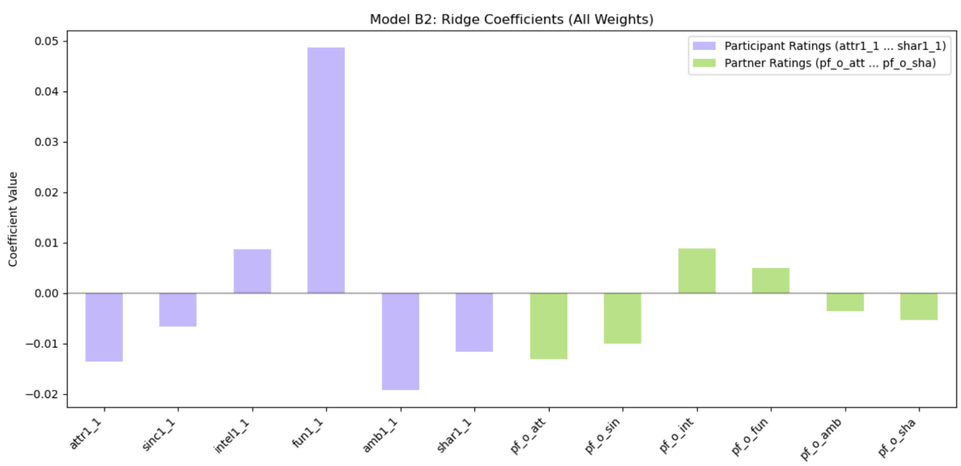


- Lasso Logistic Regression: Accuracy: 0.644, AUC = 0.703

Both models performed similarly, with Ridge slightly better than Lasso.

- Precision(match class): only about 30% of predicted matches were actually matches.
- Recall(match class): the model correctly identifies most actual matches, but at the cost of many false positives.

Model B's accuracy is overall lower than Model A's, but this is misleading because Model A is achieving its accuracy by simply predicting "no match" for everyone. Model's AUC is at chance level, implying that there are no meaningful patterns from demographics alone. Model B, however, implies that the stated dating preferences(weights) actually do provide modest predictive power and that the model is actually learning something.

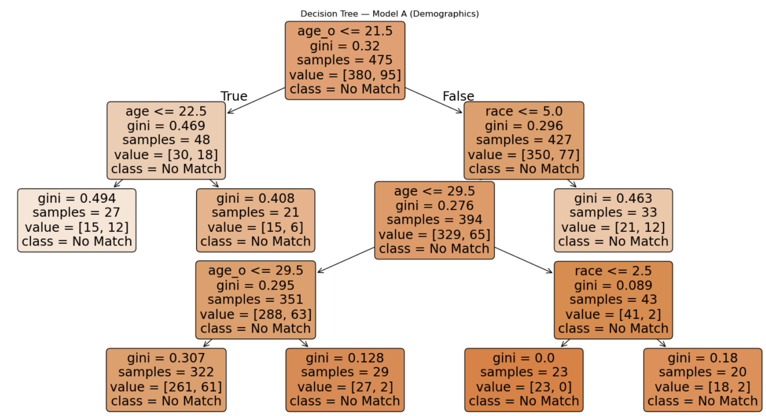


The most predictive variable was fun1_1(how much the participant values fun):

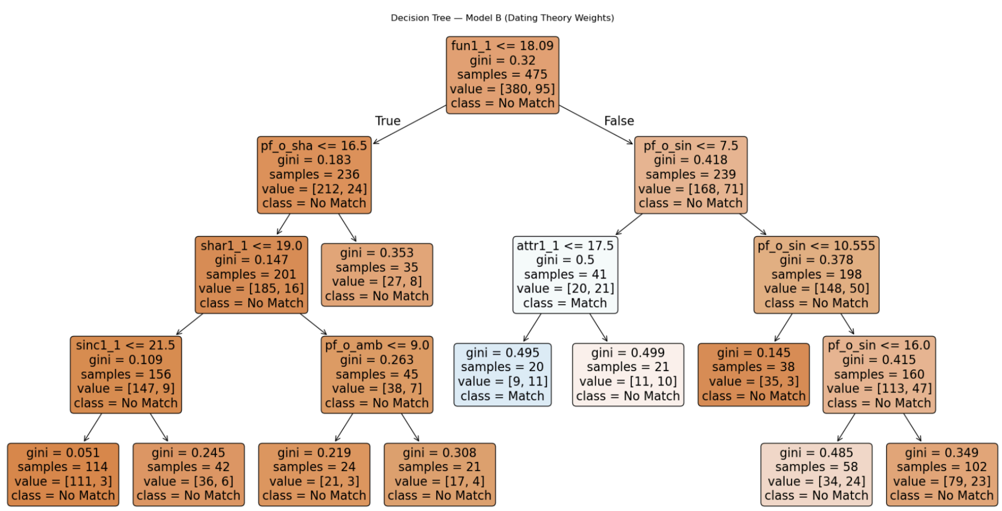
- Lasso coefficient: 0.0539
- Ridge coefficient: 0.0486

The consistency of coefficient values across Lasso and Ridge provides considerable evidence of a relationship between the attributes and match. The results also make psychological sense: fun and intelligence are positive predictors for a match. What is surprising is that an over-emphasis on attractiveness, sincerity, or ambition slightly decreases match likelihood. There are some limitations, however. All the coefficients are small, meaning that they have a relatively weak effect on whether a match actually happens.

Decision Trees

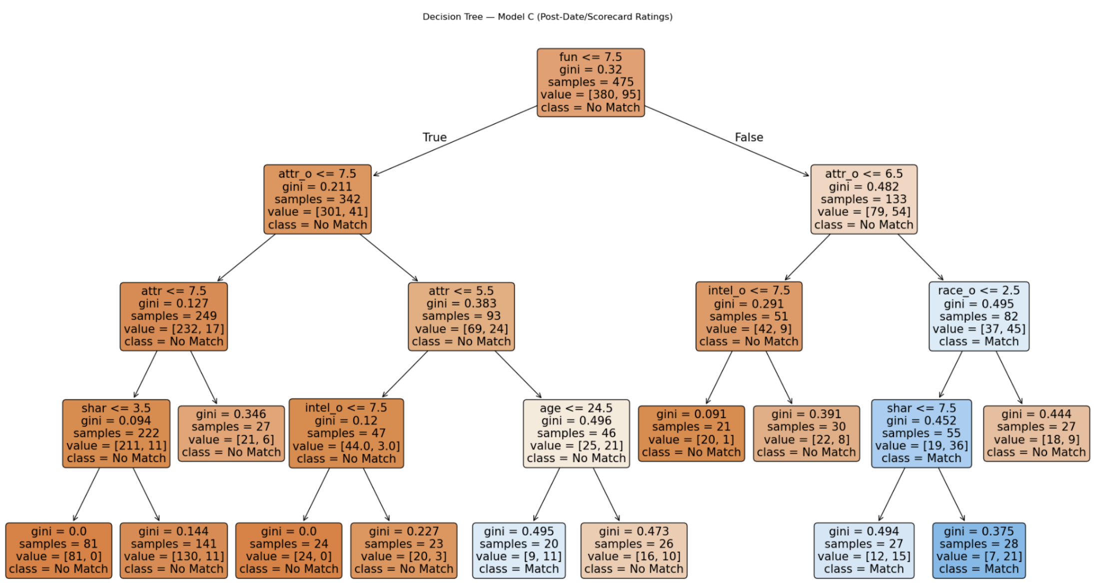


Decision Tree: Model A (Demographics)
Since the decision tree does not identify any regions that reliably predict matches, permutation importance and partial dependence were not done for this model. On their own, demographic variables may not meaningfully separate the data into regions associated with matches.



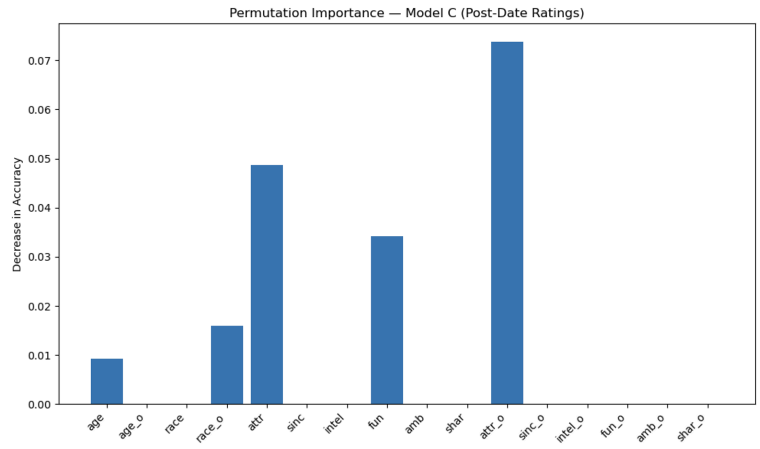
Decision Tree: Model B (Dating Theory Weights)
This decision tree suggests that matches only happen under a specific combination of preferences. Firstly, if fun is not valued strongly by the participant, then there is likely no

match. Secondly, the partner has to value sincerity relatively low for a match to potentially happen. Thirdly, attractiveness must be valued strongly by the participant for a match. It is important to note that the tree has low confidence in these distinctions, so there may be other factors that are unobserved.



Decision Tree: Model C (Post-Date/Scorecard Ratings)

This decision tree shows a few pathways to matches with limited confidence, but more than Model B's. Firstly, if participants don't strongly value fun, matches rarely happen unless both parties don't prioritize attractiveness. In these cases, matches only happen when participants are younger than 24.5 years. However, when participants do value fun, matches are more likely if partners value attractiveness. On top of that, matches are more likely if the partner identifies as White or Black (this may be due to the race imbalance of data), and also if the participant doesn't strongly prioritize shared interests. It is important to note that there is a weak separation between matches and non-matches, suggesting that these preferences only explain a small part of outcomes.



Attractiveness and fun seem to carry the most weight when it comes to match likelihood.

Discussion

Overall, the results suggest that people have only partial insight into the factors that drive their romantic decisions. Stated dating preferences, particularly valuing fun, show only modest predictive power in both logistic regression and decision tree models, implying that participants are not entirely wrong about what matters to them. Logistic regression on dating theory weights suggests that over-emphasizing attributes such as attractiveness is slightly associated with a lower likelihood of a match, though this effect is small and should be interpreted cautiously.

In contrast, Model C, which uses post-date evaluations, demonstrates substantially stronger predictive power. Permutation importance analyses show that perceived attractiveness and fun after the interaction are the most influential predictors of a match. Demographic variables, in comparison, contribute very little meaningful information.

Together, these findings suggest that while individuals possess some awareness, particularly regarding the importance of fun, there is a gap between self-reported beliefs and the factors that actually drive behavior during real interactions. This highlights the role of potential situational impressions formed during the date, which seem to outweigh stated preferences. These results may help future laboratory studies of decision-making, as well as assist the performance of dating platforms by emphasizing actual behavior over self-reported preferences.