Journal of Experimental Social Psychology 90 (2020) 103968

Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

journal homepage: www.elsevier.com/locate/jesp

Negligible evidence that people desire partners who uniquely fit their ideals☆,☆☆

Jehan Sparks[a,*], Christine Daly[b], Brian M. Wilkey[c], Daniel C. Molden[d], Eli J. Finkel[d], Paul W. Eastwick[e]

[a] *Social Cognition Center Cologne, University of Cologne, Germany*
[b] *Indiana University School of Medicine, United States of America*
[c] *University of Texas at Austin, United States of America*
[d] *Northwestern University, United States of America*
[e] *University of California, Davis, United States of America*

ARTICLE INFO

Keywords:
Ideal partner preferences
Person perception
Relationships
Predictive validity
Matching

ABSTRACT

Laypersons and scholars often presume that people positively evaluate partners who match their ideal partner preferences: If Faye prefers kindness in a partner and Sonia prefers ambition, Faye should be especially attracted to kind partners and Sonia should be especially attracted to ambitious ones. However, to date, most published tests of this idea are imprecise and permit multiple interpretations of the data. The current studies improve upon prior tests by (a) having participants self-generate the ideal attributes that matter most to them and (b) using a yoked design to isolate the predictive power of self-generated (vs. other-generated) ideal attributes. Overall, participants were more romantically interested in blind-date partners (Study 1) and acquaintances/friends/romantic partners (Study 2) to the extent that they thought those individuals possessed the ideal attributes. But the positive association of these attributes with romantic interest was identical regardless of whether the attributes represented the participant's self-generated ideals or someone else's ideals. We also used a novel coding scheme to organize participants' 1011 self-generated ideal attributes into 95 different attribute-categories; we then implemented three exclusion strategies (that differed in breadth vs. precision) using this scheme in order to maximize idiosyncratic variability between self- and other-generated ideals. All approaches revealed identical conclusions. Focused tests of ideal partner preference-matching may reveal that individual differences in ideal partner preferences poorly correspond to the attributes that uniquely inspire romantic interest.

1. Introduction

People often use rich, detailed information to describe their preferences, and idiosyncratic variability in preferences has provided inspiration for countless empirical investigations across the social sciences. One type of preference that exhibits stable, reliable individual differences is the *ideal partner preference*—the attributes (e.g., sensitive, attractive, ambitious) that inspire liking for potential or actual romantic partners (Buss, 1989; Fletcher, Simpson, Thomas, & Giles, 1999; for a review, see Eastwick, Finkel, & Simpson, 2019). Although people regularly report strong preferences for particular attributes (e.g., Faye prefers kindness, and Sonia prefers ambition), it is critical to understand whether individual differences in such reported preferences affect the way that people evaluate the potential or actual romantic partners that they encounter in real life. Prior studies offer vague and conflicting evidence on this point; the current research introduces a novel experimental approach that attempts to provide a more focused test of the predictive validity of ideal partner preference-matching.

1.1. Ideal partner preferences as a comparison standard

Classic theories on relationships assume that people possess comparison standards that they rely on when evaluating partners (Thibaut & Kelley, 1959). Building off this idea, the Ideals Standards Model posits that there are meaningful individual differences in such comparison standards, and that people positively evaluate partners to the

☆ This paper has been recommended for acceptance by Matthew Montoya.

☆☆ The authors thank Yilin Andre Wang and Matthew Baldwin for their helpful feedback at various stages of this project, and Sarah Carpenter for her research assistance. This work was supported by the National Science Foundation [NSF-0951571].

* Corresponding author.
E-mail address: jehan.sparks@gmail.com (J. Sparks).

https://doi.org/10.1016/j.jesp.2020.103968
Received 13 September 2018; Received in revised form 5 February 2020; Accepted 11 February 2020
0022-1031/

extent that those partners match their ideal partner preferences for particular traits (Fletcher et al., 1999; Fletcher & Simpson, 2000; Fletcher, Simpson, & Thomas, 2000; Simpson, Fletcher, & Campbell, 2001). That is, matches between (a) people's idiosyncratic self-reported ideals for particular traits (e.g., "I value kindness in a romantic partner") and (b) partners' levels of that same trait (e.g., "My partner is kind") should predict (c) romantic outcomes (e.g., romantic interest in the partner).

Despite the clear theoretical rationale that the match between people's idiosyncratic ideals and their perceptions of partner attributes—i.e., *ideal partner preference-matching*—should predict important outcomes, a growing body of empirical work has failed to find consistent evidence for this hypothesis in contexts where people evaluate real-life potential or actual romantic partners (e.g., Eastwick & Finkel, 2008; Eastwick, Finkel, & Eagly, 2011; Lam et al., 2016; for reviews, see Eastwick, Finkel et al., 2019; Eastwick, Luchies, Finkel, & Hunt, 2014). Although ideal partner preference-matching does predict more positive evaluations of photographs of other-sex individuals (e.g., Conroy-Beam & Buss, 2017; DeBruine et al., 2006; Eastwick & Smith, 2018; Wood & Brumbaugh, 2009), as soon as two partners interact face-to-face, such matching has weak or nonexistent effects on partner evaluations. These findings could be interpreted as indicating that individual differences in ideal partner preferences do not reflect the attributes that truly inspire one's romantic desire in real-life potential partners.

1.2. Optimizing tests of the predictive validity of ideal partner preference-matching

The studies documenting a lack of predictive power in face-to-face contexts have generally operationalized an "ideal match" in terms of the absolute levels of ratings on a relatively large set of researcher-generated traits. This operationalization addresses the conceptual question, "If participants report especially high ideals on a particular trait, do they report more positive romantic outcomes if their partner has that trait?" For example, if participants highly value kindness in a partner, do they report more positive romantic outcomes to the extent that the partner is kind? In studies that use this type of *level metric* (see Eastwick et al., 2014), predictive validity is assessed trait-by-trait by examining whether the statistical interaction between (a) an ideal trait (e.g., kindness) and (b) partners' perceived traits positively predicts (c) romantic outcomes (for a related approach, see Wood & Brumbaugh, 2009).

This approach may be too conservative. One explanation for the poor predictive validity of these tests is that participants in real life may focus on only a few personally important traits rather than the full suite of traits presented to them by a researcher. Real-life partners are complex and possess many traits that may be difficult to evaluate when bundled together (Li, Bailey, Kenrick, & Linsenmeier, 2002): For example, a potential mate may be high in attractiveness but low in honesty and moderately funny. It would require immense cognitive effort to compute and match preferences across all traits in the search for an ideal partner. Thus, real-life romantic evaluations necessarily involve simplifications and tradeoffs, where some ideals will be met at the expense of others (Conroy-Beam, Goetz, & Buss, 2016). Although researchers might have assessed participants' ideals for 15 (or more) traits, each participant in reality may only have considered a subset of those ideals (e.g., Faye considers the attributes kind, intelligent, and attractive, whereas Sonia considers ambitious, generous, and adventurous). If different participants consider different ideals, the overall effect size for a given Ideal × Trait interaction will be quite small even if some participants are drawing on their ideals for that trait. Thus, poor predictive validity in face-to-face contexts could simply be an artifact of a study design that relies on large numbers of ideals—a design which has caused scholars to overlook the fact that participants were indeed considering the match between (some of) their ideals and a partner's traits.

There are alternative approaches to the level metric that have demonstrated support for the predictive validity of ideal partner preference-matching, but these approaches also have major limitations. Specifically, alternative tests of predictive validity (e.g., approaches using ideal trait pattern-matching, Fletcher et al., 2000, 1999; approaches using a Euclidean distance metric, Conroy-Beam et al., 2016) are confounded with the desirability of the traits themselves, a problem called the *normative-desirability confound* (Rogers, Wood, & Furr, 2018; Wood & Furr, 2016; Wood, Lowman, Harms, & Roberts, 2019). In other words, the predictive power of these approaches likely reflects the fact that people tend to report more positive romantic outcomes if a partner has more positive traits, rather than reflecting something about the match between ideals and partner traits. It is possible to algebraically remove the normative desirability confound when testing the predictive validity of preference-matching (e.g., Wood et al., 2019), but in the ideal partner preferences domain, this mathematical procedure is not yet common. In one study that used it (Lam et al., 2016, Study 4), researchers found small and nonsignificant effect sizes for ideal partner preference-matching in a U.S. sample (i.e., $\beta \approx 0.05$) but a modestly sized effect in a Taiwanese population (i.e., $\beta \approx 0.22$). Also, Eastwick, Finkel, & Simpson (2019) reported a reanalysis of an earlier study (Eastwick et al., 2011, Study 3) in which the effect of ideal partner preference-matching dropped from $r = .19$ to $r = -.04$ after addressing the normative desirability confound. In short, the presence of normative desirability may render some tests of preference-matching too liberal; an experimental approach that reduces the influence of normative desirability without the need to make these algebraic corrections would be a valuable addition to the literature.

In the current research, we developed a procedure to assess the predictive validity of ideal partner preference-matching that addresses the limitations of previous studies. First, if we accept the premise that people face tradeoffs when evaluating potential partners (Conroy-Beam et al., 2016), it follows that one way of maximizing our chances at finding predictive validity evidence is to ask people to consider only their most important ideals—the attributes that they personally view as critical necessities to be prioritized first and foremost (Li et al., 2002; Li & Kenrick, 2006). Thus, to ensure that our test only weighs ideals that truly matter to people (rather than a large set of researcher-generated ideals), participants nominated their three most important ideal attributes in a partner. The proposition that a personal top-three set of ideals might prove to have especially strong predictive power follows from other studies of mate preferences positing that participants commonly express idiosyncratic priorities for certain attributes over all others in the form of "trade-offs" or possess personal "dealbreakers" and "dealmakers" (Fletcher, Tither, O'Loughlin, Friesen, & Overall, 2004; Joel, Teper, & MacDonald, 2014; Jonason, Garcia, Webster, Li, & Fisher, 2015; Li et al., 2002; Li & Kenrick, 2006).

Second, to address the normative-desirability confound, participants also rated potential relationship partners on the three most important ideal attributes generated by another random same-sex participant using a yoked design. Yoked designs are within-subjects experimental manipulations that prove especially useful when researchers are interested in testing the idiosyncratic priority of a small set of items selected from a larger population of items (e.g., Przybylinski & Andersen, 2015). In this case, the predictive validity of ideal partner preference-matching would be supported if the association of participants' perceptions of a partner's attributes with romantic interest is stronger for self-generated rather than other-generated ideal attributes.

1.3. The current research

In the current studies, participants first nominated their top-three ideals. They then went on a blind-date (Study 1) or nominated five targets of their preferred sex with whom they were already acquainted, including a current romantic partner if they had one (Study 2). They rated the extent to which these targets possessed the attributes they nominated as ideal ("self-generated" ideal attributes) and the attributes nominated by a same-sex yoked partner as ideal ("other-generated" ideal attributes). Finally, participants reported on their romantic interest in each target as the dependent measure. In all studies, we report all measures, manipulations and exclusions either in the main text or in Supplemental materials.

2. Study 1

Study 1 tested the unique role of ideal partner preference-matching using a blind-date paradigm with other-sex dyads. Participants listed their three most important ideal attributes in a romantic partner, which could have been traits or something more specific (e.g., behaviors, see Study 2). Participants then went on a date, and afterwards, rated the extent to which their blind-date partner possessed the three self-generated ideal attributes and three other-generated ideal attributes (nominated by one other unknown participant in the study, see below). Finally, they reported their romantic interest in their blind-date partner.

If ideal partner preference-matching uniquely predicts romantic interest, then matches with participants' self-nominated, most important ideals should predict romantic interest more strongly than matches with the most important ideals nominated by someone else. To illustrate: Suppose that Faye nominates *kind* as an ideal attribute, whereas her yoked partner Sonia nominates *ambitious* as an ideal attribute. Because kind and ambitious are both positive traits, it is likely that both participants will report more romantic interest to the extent that they perceive the partners to be kind and ambitious. However, if people's ideal partner preferences are uniquely predictive of relationship outcomes, then a partner's kindness should be more strongly associated with romantic interest for Faye (vs. Sonia), and a partner's ambition should be more strongly associated with romantic interest for Sonia (vs. Faye).

2.1. Method

2.1.1. Participants

Participants were 138 single, heterosexual undergraduates from Northwestern University (69 women, 69 men) who were recruited to participate in a study on heterosexual relationship formation by "going on a blind-date for science." Each participant received $40 for participation, with $10 given in advance to spend on the date. In Study 1, participants were 19.2 years old on average (SD = 1.1 years). Approximately 68.1% of participants reported that they were Caucasian, 2.9% were African American, 15.2% were East Asian, 14.5% were South Asian, 3.6% were Hispanic, 1.4% were Middle Eastern, and 1.4% were Native American (participants were able to select all races that applied). We collected as many participants as possible over two academic quarters; we did not conduct analyses until data collection was complete.

The research team matched other-sex dyads for the blind-dates using information from an initial screening survey to ensure a woman was randomly selected for each man that was (a) the same race, (b) the same height as or shorter, (c) either the same year in college or within one year, and (d) available for the blind-date session. Of the 69 dyads matched, one reported that they did not spend their date together, and three reported knowing each other (i.e., they provided a response of 4 or greater on a 1 [*not at all*] to 7 [*extremely well*] scale assessing how well they had known their partner before the session). Therefore, 65 dyads were retained for analysis. Of these N = 130 participants, 2 were excluded because they (erroneously) listed negative rather than positive ideal attributes (see below), leaving a total usable N = 128. (Hypothesis tests in Supplemental materials reveal identical conclusions if the three dyads who reported knowing each other are included.)

A sensitivity power analysis in this context calculates a self-generated-attribute/romantic-interest association, given a sample size, an other-generated-attribute/romantic-interest association, and a self-generated-attribute/other-generated-attribute association. We used the data we ultimately obtained in Study 1 for the other-generated-attribute/romantic-interest association and the self-generated-attribute/other-generated-attribute association as inputs to solve for the self-generated-attribute/romantic-interest association. Thus, in Study 1, assuming an other-generated-attribute/romantic-interest association of $\beta = 0.36$ and a self-generated-attribute/other-generated-attribute association of $\beta = 0.42$, $N = 125$ gives us 80% power to detect a self-generated-attribute/romantic-interest association of $\beta = 0.58$ (and thus, the difference between 0.58 and 0.36 is $\beta_{\text{difference}} = 0.22$, which is a small-to-medium-sized effect, Cohen, 1992). All power analyses were conducted in G*Power (Faul, Erdfelder, Buchner, & Lang, 2009), and like any power analysis, they reflect an educated guess; simulations might have produced different estimates.

2.1.2. Procedure

2.1.2.1. Part 1: pre-date questionnaires. Participants first reported demographic information to determine eligibility and match the blind-date dyads. Approximately one week prior to the date, participants completed a 60-min online *pre-date questionnaire*.

Prior to Part 2, pairs of same-sex participants were randomly yoked to one another to create the self-generated and other-generated ideal attribute variables. The yoking was run in batches as eligible participants were recruited. Two larger batches of participants were yoked at the beginning of each academic term; smaller batches were then yoked weekly as additional participants were recruited.

2.1.2.2. Part 2: blind-date session. A research assistant escorted the male and female blind-date dyad members to a waiting room where they were introduced to each other for the first time. After a three-minute getting-acquainted period, the research assistant explained the conditions of the blind-date and gave each participant $10. The blind-date lasted 1 hour, and final payment was contingent upon the dyad's timely return to the laboratory. The dyad was not otherwise restricted in terms of location or activity, and they could choose to spend any portion of their $10 payment on the date or save it all. Most participants went to eat at a dining hall on campus, others went to eat off campus or got a coffee, and others simply walked around campus and chatted.

On returning from their date, the research assistant greeted the dyad and led them to separate cubicles where they each completed a 30-min *post-date questionnaire*. After completing the questionnaire, participants were debriefed and given the remainder of their payment. All variables collected in the questionnaires are included ("Study 1 list of all variables collected") at the following link: https://osf.io/hdpyw/?view_only=462da328a82049bcb9bc14bf41e4dbf5.

2.1.3. Materials

2.1.3.1. Soliciting ideal attributes. Participants' self-generated ideal attributes were recorded on a free response item in the pre-date questionnaire:

> Please tell us what you think are important qualities you consider when deciding whether to pursue a romantic relationship with someone … You can write single words, like personality traits, or brief descriptions of a specific quality or behavior. Whatever is most important to you. Please list three of these qualities.

2.1.3.2. Coding ideal attributes into attribute-categories. Across all studies reported in the present research, our nomination procedures led participants to generate 1011 unique ideals (see "Codebook": https://osf.io/hdpyw/?view_only=462da328a82049bcb9bc14bf41e4dbf5). The first and last authors coded ideals into attribute-categories based on the 66 categories used in Fletcher et al.'s (1999) Tables 1 and 2, which is, to our knowledge, the only other effort to categorize naïve

Table 1
Summary of attribute exclusion approaches: Studies 1 and 2.

	Exclusion scheme	Exclusion rule	Rule breadth	Exclusion rates			Rationale	Method of combining traits	Origin of rationale	Support for predictive validity	
				Ideals	Participants	Attributes per participant				β_{dif} *M*	β_{dif} range
1	Attribute-level (primary analyses)	2 ideals match one of 95 attribute-categories	Narrow	Study 1: 16.8% Study 2: 15.8%	Study 1: 2.3% Study 2: 0.5%	Study 1: 5.04 Study 2: 5.04	Adopts the high fidelity of traits as defined in Fletcher et al. (1999, Tables 1 and 2)	Trained coder, κ = .91	Requested in round 1 reviews	.03	.00-.10
2	Synonym-level (supplemental materials)	2 ideals match 1 of 10 "groups" or leftover 53 attribute-categories	Intermediate	Study 1: 26.2% Study 2: 30.6%	Study 1: 3.1% Study 2: 4.5%	Study 1: 4.47 Study 2: 4.18	Adopts an intermediate level of precision by combining across top synonyms only	Top synonyms of the 95 attributes at thesaurus.com	Requested in round 2 reviews	.04	-.01-.08
3	Factor-level (supplemental materials)	2 ideals match 1 of 3 factors or leftover 40 attribute-categories	Broad	Study 1: 58.1% Study 2: 71.0%	Study 1: 29.7% Study 2: 33.1%	Study 1: 2.91 Study 2: 2.56	Adopts the broad bandwidth implied by participants' typical ratings of ideals and attributes	Factor analysis of ideals reported in Fletcher et al. (1999)	Original analysis (*a priori*) in initial submission	.04	-.05-.14

Note. Summary of the three different approaches to determining which, if any, of the six attribute ratings (three for self-generated, three for other-generated) should be excluded for a given participant. Each approach relies on the coding procedure described in the main text, but differs in the extent to which they determine exclusions precisely (attribute-level) versus broadly (factor-level). β_{dif} refers to the difference between the self-generated-attribute/romantic interest beta and the other-generated-attribute/romantic interest beta.

participants' spontaneously generated ideals. Given that the number of participants (and hence the number of ideals) in the present work is considerably larger than Fletcher et al. (1999), we observed a wider range of ideals, and thus we added 29 new attribute-categories to fully capture the diverse array of ideals generated by our participants. In total, ideals were categorized into 95 attribute-categories. Appendix A contains the percentage of total ideals nominated at pretest by attribute-category, both by study and overall. These percentages correlate with Fletcher et al.'s (1999) rate of mention of attribute-categories at $r = 0.68$, suggesting these are the same constructs as reported in past studies. In addition, we had an independent coder categorize the 1011 ideals into attribute-categories using a codebook containing our list of 95 attribute-categories and the participant-generated ideal that was the closest synonym to each attribute-category (see column G in "Codebook": https://osf.io/hdpyw/?view_only=462da328a82049bcb9bc14bf41e4dbf5). Results revealed strong agreement between the coder's categorizations and the coding scheme produced by the first and last authors, $\kappa = 0.91, p < .001$. The fact that kappa is even higher ($\kappa = 0.93, p < .001$) if each attribute-category is weighed by how commonly it is nominated suggests that the few disagreements tended to emerge for the rare attribute-categories. Indeed, the rare attribute-categories in which the first and last author disagreed with the coder would only have affected 2.9% of the exclusion decisions. (Exclusion decisions are discussed in more detail below.)

Participants nearly always listed ideals that were framed positively—that is, qualities they wanted an ideal partner to possess (e.g., considerate) rather than not possess (e.g., inconsiderate). In our analyses, we excluded the two participants' whose responses were framed negatively (i.e., inconsiderate, smoking). We made these exclusions because the association between the ideal partner attribute and romantic interest should be positive for the positively framed ideal attributes but negative for the negatively framed ideal attributes. Thus, the magnitude of the attribute/romantic-interest association loses its meaning if it is built from an aggregation of positive and negative ideal attributes.

Table 2
Descriptive statistics of primary variables: Study 1.

Variable	Mean	SD
Romantic interest (dependent variable)	3.16	1.30
Self-generated ideal attribute ratings (independent variable for self-generated regression)	5.04	1.17
Other-generated ideal attribute ratings (independent variable for other-generated regression)	5.12	0.98

Note. Means and standard deviations from the unstandardized variables used in the self-generated regression model and the other-generated regression model. All variables were measured on a scale from 1 (*strongly disagree)* to 7 (*strongly agree*).

2.1.3.3. Rating target attributes. In the post-date questionnaire, participants were asked to rate on a scale from 1 (*strongly disagree)* to 7 (*strongly agree*) the extent to which they believed six attributes (3 self-generated ideals, 3 other-generated ideals) were characteristic of their blind-date partner.

2.1.3.3.1. Self-generated ideal attribute ratings. Participants rated their partner on each of the three ideal attributes they listed on the pre-date questionnaire (e.g., Faye would rate her partner on her three nominated ideal attributes: kind, intelligent, and attractive). These ratings were then averaged to create a composite score representing the extent to which participants perceived that their blind-date possessed the three self-generated ideal attributes ($\alpha = .60$). Given the relatively low reliability, we include ideal attribute-by-ideal attribute analyses in Supplemental materials; hypothesis tests reveal identical conclusions.

2.1.3.3.2. Other-generated ideal attribute ratings. In addition, participants rated their blind-date on the three ideal attributes that another randomly selected, same-sex participant (i.e., their yoked partner) listed on the pre-date questionnaire (e.g., Faye would rate her partner on Sonia's three nominated ideal attributes: ambitious, generous, and adventurous). These ratings were then averaged to create a composite score representing the extent to which participants perceived that their blind-date partner possessed the other-generated ideal attributes (α = .42). Again, given the low reliability, we include ideal attribute-by-ideal attribute analyses in Supplemental materials; hypothesis tests reveal identical conclusions.

2.1.3.4. Rating romantic interest. Participants reported their romantic interest in their blind-date by rating them on four items ("I really like this person," "I am sexually attracted to this person," "I am romantically interested in this person," and "This person and I have a passionate connection") on a 1 (*strongly disagree)* to 7 (*strongly agree*) scale (α = .84).

2.1.3.5. Exclusions. The yoked design requires that the participant and his/her yoked partner idiosyncratically vary to some extent in their three nominated ideals. That is, if self- and other-generated ideals are "too similar," such similarity limits the opportunity to find a difference in the self-generated ideal attribute versus other-generated ideal attribute associations with romantic interest. Any operationalization of "too similar" will necessarily confront the classic bandwidth-fidelity tradeoff in the way traits are categorized (Cronbach & Gleser, 1957; John, Hampson, & Goldberg, 1991). For example, *trustworthy* and *stable* are semantically distinct attributes with different definitions, and a high-fidelity approach (like Fletcher et al., 1999, Tables 1 and 2) would indeed treat them as two different ideal partner preferences. But they are synonyms of each other (www.thesaurus.com), and perhaps a typical participant would assume that *trustworthy* subsumes *stable*, and vice versa. In principle, one could further advance this similarity argument to apply to attributes that are not synonyms but nevertheless tend to be associated with each other. When participants rate actual ideals and attributes of partners, they perceive that some traits go with other traits (Fletcher, Kerr, Li, & Valentine, 2014; Overall, Fletcher, & Simpson, 2006); people think that a *trustworthy* and *stable* person is also a *kind* person, even though *kind* is not a synonym of *trustworthy* or *stable*. This observation suggests the utility of a broad bandwidth approach to exclusions that incorporates people's beliefs about the way traits covary with other traits.

To address these issues, our manuscript adopts three different approaches to determining which, if any, of the six attribute ratings (three for self-generated, three for other-generated) should be excluded for a given participant. All approaches capitalize on the coding procedure described above, and they differ in the extent to which they determine exclusions precisely versus broadly (Table 1).

First, all primary analyses in the main text use the narrowest approach: an attribute-category exclusion approach that sticks to the 95 attribute-categories in Appendix A. In this approach, we eliminated any self-generated and other-generated ideal attribute ratings that came from the same attribute-category when creating the self-generated and other-generated ideal attribute averages for each participant (e.g., if a participant generated *smart* and her yoked partner generated *intelligence*, these ideal attribute ratings were excluded because both ideals came from the attribute-category Intelligent; the self-generated and other-generated ideal measures were then created from the remaining two unique ideals instead of all three). The top row in Table 1 describes the exclusion rates based on this approach.

Second, a moderately precise (i.e., synonym-level) approach is reported in Supplemental Materials. This approach excludes attributes based on 10 groups determined by common synonyms taken from thesaurus.com; these 10 synonym groups encompass 42 of the 95 attributes in Appendix A. This approach would eliminate the participant's ratings on the attributes *stable* and *trustworthy*, for example, because they belong to the same attribute group. If the ideals did not fit into one of the 10 groups, they were excluded if they matched at the attribute-category level for one of the remaining 53 attributes (as in our primary, attribute-category exclusion approach). The middle row in Table 1 describes the exclusion rates based on this approach.

A third, broad approach is also reported in Supplemental Materials. This approach draws from the three Fletcher et al. (1999) attribute factors: (a) Warmth-Trustworthiness/Intimacy-Loyalty, (b) Vitality-Attractiveness/Passion, and (c) Status-Resources. According to Fletcher et al. (1999), each of these three categories "reflects a possible 'solution' to a specific barrier to successful reproduction" (p. 87) that characterized humans' ancestral past. In other words, even though attributes like *attractive* and *adventurous* are not colloquially synonymous, they both belong in the factor "Vitality-Attractiveness" because they can be used to identify a partner who is healthy and reproductively viable. Furthermore, when participants rate other people on attributes like those in Appendix A, the ratings generally fit this three-factor solution (e.g., Fletcher et al., 2014; Overall et al., 2006). Thus, this approach excludes ideals if they matched on one of the three factors from Fletcher et al. (1999), and if the ideals did not fit into one of the three factors, they were excluded if they matched at the attribute-category level for one of the remaining 40 attributes (as in our primary, attribute-category exclusion approach). The third row in Table 1 describes the exclusion rates based on this approach.

2.2. Results

Because our design randomly assigned participants to both a blind-date dyad and a yoked-pair dyad, our data were structured such that each row represented an individual participant with separate variables that kept track of which participants constituted a yoked-pair and which participants went on a blind-date. The partially hierarchical nature of the data led us to test different multilevel models using the lme4 package (Bates, Maechler, Bolker, & Walker, 2015) in the R environment (R Core Team, 2018). These included a cross-classified random intercepts, fixed slopes model that nested participants (level 1) within both a blind-date dyad (level 2) and a yoked-pair dyad (also level 2), and a simpler random intercept, fixed slopes model with participants (level 1) nested within a yoked-pair dyad (level 2). Given that the variance of our random effects were very small and even zero in some cases (especially for blind-date dyad), we also tested a simple linear regression model. The linear regression model specification provided the best fit to our data in terms of the lowest Bayesian Information Criteria (BIC) values, and thus we present results from this model in our primary analyses; hypothesis tests reveal identical conclusions regardless of model chosen (see Supplemental materials for summary of results across models tested for primary analyses, and see link to "JESP Models Tested Studies 1 and 2.html" document with R code and results from all models tested: https://osf.io/7fq9n/?view_only=336e96f6730b480182be7afe49383d31).

We standardized all variables, and then ran two separate models: (a) a self-generated model that predicted participants' romantic interest in their blind-date from the extent to which their self-generated ideal attributes were rated as characteristic of their blind-date, and (b) an other-generated model that predicted participants' romantic interest in their blind-date from the extent to which their other-generated ideal attributes were rated as characteristic of their blind-date.

The self-generated ideal attribute ratings positively predicted romantic interest, $\beta = 0.46$, $SE = 0.08$, $p < .001$, as did the other-generated ideal attribute ratings, $\beta = 0.36$, $SE = 0.08$, $p < .001$ (see Table 2 for descriptive statistics of all variables used in regression models). However, and of primary theoretical interest, these two associations ($\beta = 0.46$ and $\beta = 0.36$) were similar in magnitude. To test the significance of the difference between these two regression coefficients generated by the same set of participants (i.e., two dependent associations), we used Lee and Preacher's (2013) web utility, which implements Steiger's (1980) formulas for comparing such coefficients. Importantly, this statistical test controls for the correlation between self-generated and other-generated ideal attribute ratings, $\beta = 0.42$. The two associations (i.e., $\beta = 0.46$ and $\beta = 0.36$) did not significantly differ from each other, $z = 1.16$, $p = .245$ (see Fig. 1). Furthermore, sex did not moderate the association between self-generated ideal attribute

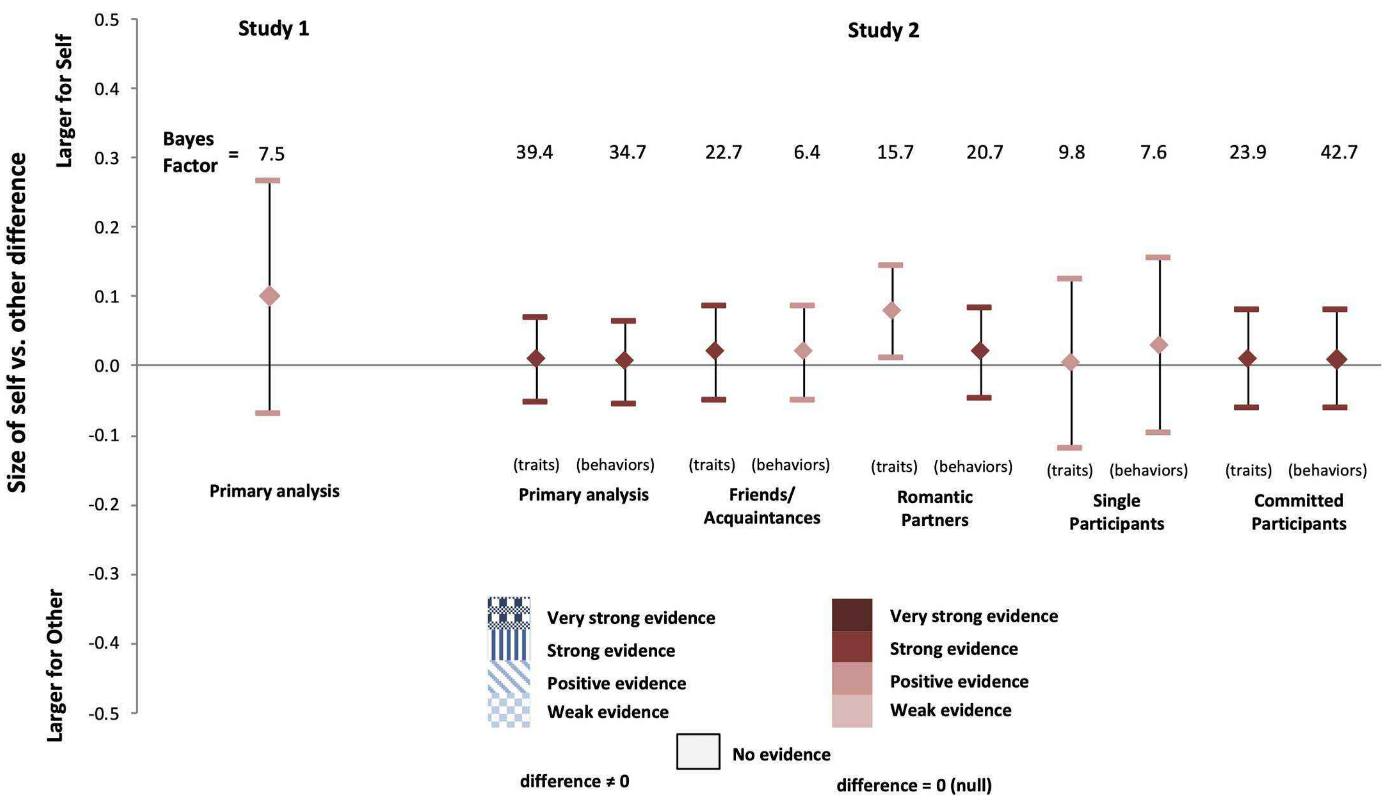

Fig. 1. Effect sizes, 95% confidence intervals, and Bayes factors indicating support for H_0: Studies 1 and 2.
Note: Plot showing the difference in the strength of association between (a) self-generated ideal attribute ratings and romantic interest and (b) other-generated ideal attribute ratings and romantic interest in each sample reported in Study 1 and Study 2. The size of the difference is close to zero, indicating that self-generated ideal attribute ratings do not predict romantic interest more strongly than other-generated ideal attribute ratings. Error bars indicate 95% confidence intervals as calculated from regression (Study 1) or multilevel regression (Study 2). Bayes factors (BFs) reflect the strength of the evidence (i.e., 55% "strong" and 45% "positive") for the null hypothesis based on SEM as described by Wagenmakers (2007, Table 3). In cases of disagreement between 95% confidence intervals and SEM (e.g., romantic partners, traits), quantitatively minded scholars generally consider SEM results to be more accurate (Kline, 2005; Ledgerwood & Shrout, 2011).

ratings and romantic interest, $\beta = -0.06$, $SE = 0.15$, $p = .690$, and it did not moderate the association between other-generated ideal attribute ratings and romantic interest, $\beta = 0.06$, $SE = 0.16$, $p = .718$. Thus, participants expressed more romantic interest in their blind-date partners to the extent that they perceived partners to possess positive attributes; however, it did not matter whether those attributes matched the three ideals they desired most in a partner or the three ideals that another random participant desired most in a partner. That is, there was no evidence for the unique role of ideal partner preference-matching in predicting romantic interest.

2.3. Discussion

Study 1 did not support the unique predictive validity of ideal partner preferences for interest in a blind-date partner using a yoked design. To the extent that the blind-date partner possessed attributes that the participant nominated as an ideal, the participant was more romantically interested in the partner. But this association appeared to be equally strong for other attributes that a *different* participant nominated as ideal. In other words, participants liked blind-date partners to the extent that they thought those partners had "dealmaker" attributes (Jonason et al., 2015), but there was no boost in predictive power if the attributes represented ideals held by the participants themselves or ideals held by someone else.

Importantly, however, Study 1 was underpowered to detect a small effect size of our manipulation; note, for example, that the 95% confidence interval for the difference between the two associations contains the null (i.e., $\beta_{difference} = 0.00$), the effect size proposed in our sensitivity power analysis (i.e., $\beta_{difference} = 0.22$), and a negative effect size as low as $\beta_{difference} = -0.07$. Given our results, it seems unlikely that there is a medium-sized or large effect of ideal partner preference matching using this manipulation; nevertheless, we recruited a much larger sample in Study 2 to achieve more precision in our effect size estimates.

3. Study 2

Study 2 had two aims. First, it applied our yoking method to a different paradigm in which participants reported on targets of their preferred sex they had previously met (including current romantic partners). Thus, this paradigm enabled us to examine whether the predictive validity of ideal partner preference-matching varies depending on the nature of the relationship between the target and the participant (i.e., acquaintance/friend vs. current romantic partner).

Second, it explicitly separated ideals for *traits* from ideals for *behaviors*. Historically, the ideals literature has defined ideal partner preferences in terms of traits, but there are reasons to believe that the literature could benefit from expanding the definition of these preferences to include behaviors. After all, relationship researchers have long been interested in the importance of partner behaviors in a variety of domains, including relationship formation, maintenance, resolution, and partner responsiveness (e.g., Finkel, Simpson, & Eastwick, 2017; Gottman, 1998; Kelley et al., 1983; McNulty & Karney, 2004; McNulty & Russell, 2010). For example, couples develop particular expectations and norms with respect to the behaviors that communicate support and enthusiasm (Gable, Gosnell, Maisel, & Strachman, 2012; Lakey & Orehek, 2011). Thus, it would be consistent with the ideal standards model if people evaluated partners against behavioral ideals in a similar manner that they evaluate partners against trait-based ideals. In other words, people should judge their relationships more positively not only when the traits of their current partner more closely match their ideal traits, but also when their current partner performs behaviors that align more closely with their ideal partner behaviors (McNulty & Karney, 2004). Moreover, since ideal behaviors may be more idiosyncratic than ideal traits, it seems plausible that ideal partner preference-matching for behaviors may be more likely to display unique predictive validity than ideal partner preference-matching for traits. Thus, Study 2 tested whether matching for behaviors vs. traits might have divergent predictive power for romantic interest.

Participants in Study 2 were instructed to report both their top three most important *traits* and their top three most important *behavioral tendencies* in an ideal romantic partner. They then reported the extent to which five targets (i.e., individuals of their romantically preferred sex whom they knew) exhibited a set of self-generated and other-generated ideal traits and behaviors. Finally, they reported on their romantic interest and relationship with each target.

3.1. Method

3.1.1. Participants

Participants were 1438 Amazon Mechanical Turk workers who were recruited to participate in Part 1. Of these participants, 909 expressed interest in completing a "follow-up survey" and provided a valid email address, and were thus eligible to complete Part 2. Additionally, only those participants who followed instructions for what qualified as an ideal trait and an ideal behavior were considered eligible. Our goal was to collect as many participants as possible in Part 1 until at least 900 participants expressed interest in completing Part 2 to allow us to reach our power analysis goal (see below) for participants reporting on romantic partners. Of the 909 eligible participants, 311 did not complete Part 2, leaving a final sample of $N = 598$ participants (425 women, 173 men). Participants received an initial payment of $0.25 for completing Part 1 of the study and an additional $0.75 for completing Part 2. In Study 2, participants were 32.4 years old on average ($SD = 10.6$ years). Approximately 74.1% of participants reported that they were Caucasian, 7.9% were African American, 5.7% were Asian, 5.2% were Hispanic, 0.2% were Native American, 5.2% were Multiracial, and 1.8% either chose other, not to disclose, or N/A (participants were able to select all races that applied). When Part 1 of the study was conducted, approximately 74.1% of participants reported being in a romantic relationship and 25.9% of participants reported being single. This relationship status variable was used to determine who is committed versus single for our relationship status analyses.

3.1.2. Procedure and materials

Study 2 consisted of two online questionnaires (i.e., Part 1 and Part 2).

3.1.2.1. Part 1. Participants began by listing their ideal partner traits and behaviors. To directly address the issue in Study 1 whereby two participants generated negative attributes, we added more details to the instructions to ensure that participants framed their attributes positively. For ideal partner traits, the instructions emphasized:

> We'd like you to list three traits/attributes that you want your ideal partner to have. Please word each trait/attribute such that it describes something you want them to have. For instance, do not write "uncool"; write "cool."

For ideal partner behaviors, the instructions emphasized:

> We'd like you to list three specific behaviors that you want your ideal partner to perform regularly. Please word each specific behavior such that is describes something you want them to perform. For instance, do not write "sets fire to my home"; write "does not set fire to my home."

Participants then provided race/ethnicity, age, and relationship status information.

We examined participants' self-generated ideal traits and behaviors prior to Part 2 to screen out participants who did not follow instructions (e.g., by generating inappropriate behavior examples). The remaining participants were randomly yoked together in same-sex pairs to create the questionnaires for Part 2.

3.1.2.2. Part 2

3.1.2.2.1. Target nominations. In Part 2, participants were asked to provide the first and last initial of five individuals whom they know personally. They were instructed to choose individuals of their romantically preferred sex, not related to them, around the same age as them, and whom they had met in person. Participants who were in a romantic relationship were instructed to list their current romantic partner as one of the five individuals. Four participants failed to follow instructions and listed the same individual multiple times. For these participants, we retained data from only the first listed target. Of the 2974 targets, 8.5% were spouses or fiancés, 8.4% were boyfriends/girlfriends, 8.5% were casual romantic/sexual partners, 51.4% were friends, 13.1% were colleagues or co-workers, 8.3% were acquaintances, 0.6% were strangers or people whom the participant had just met, and 1.2% were unreported. Partners were coded as "romantic partners" if the participant selected the "spouses or fiancés" or "boyfriends/girlfriends" categories; otherwise, the partner was coded as "not a romantic partner."

3.1.2.2.2. Rating target traits and behaviors. Similar to Study 1, participants were asked to rate on a scale from 1 (*extremely uncharacteristic*) to 11 (*extremely characteristic*) the extent to which they believed the self-nominated and other-nominated ideal traits and behaviors were characteristic of each of their five targets. That is, participants rated targets on each of the three ideal traits and behaviors they listed in Part 1 (i.e., self-generated ideal traits and self-generated ideal behaviors); ratings were averaged to create two composite scores representing the extent to which their self-generated ideal traits ($\alpha = .80$) and self-generated ideal behaviors ($\alpha = .76$) described each target. Also, participants rated targets on the three ideal traits and behaviors that their yoked partner listed in Part 1 (i.e., other-generated ideal traits and other-generated ideal behaviors). Participants' ratings on each of their yoked partner's ideals were averaged to create two composite scores representing the extent to which these other-generated ideal traits ($\alpha = .76$) and other-generated ideal behaviors ($\alpha = .77$) described each target.

3.1.2.2.3. Rating romantic interest. Participants reported their romantic interest in each of their five nominated targets on six separate statements ("I am romantically interested in ___," "I feel a great deal of sexual desire for ___," "___ is the only person I want to be romantically involved with," "___ always seems to be on my mind," "___ is very much my ideal romantic partner," and "___ and I have a lot in common") on a 1 (*strongly disagree*) to 11 (*strongly agree*) scale ($\alpha = .94$).

3.1.3. Exclusions

Our primary trait analyses reported in the main text reflect the same attribute-category exclusion approach as in Study 1; see Table 1 for exclusion rates. Results for the synonym-level, factor-level, and no-exclusion approaches are included in the Supplemental materials. All hypothesis tests revealed identical conclusions using all four approaches except for one: The one significant finding that emerged in the primary trait analyses (i.e., romantic partners, $p = .018$; see below) is $p = .031$ in the synonym-level approach (see Fig. S3 in Supplemental materials), but greater than $p = .05$ in the factor-level and no-exclusion approaches (see Figs. S4, and S5 in Supplemental materials).

For our behavior analyses, ideal behavior ratings were excluded only if ideals were worded identically (rare) or matched another behavior from the same behavior category (e.g., "housework/cleaning," "spending time together;" see details of duplicate behavior categories in Supplemental materials). This procedure allowed us to eliminate any self-generated and other-generated ideal behavior ratings that came from the same category when creating the self-generated and other-generated ideal behavior averages for each participant; 2.7% of ideal behavior ratings and no participants were eliminated using this procedure. Results for ideal behaviors using a no-exclusion approach are included in the Supplemental materials; all hypothesis tests revealed identical conclusions (see Fig. S5 in Supplemental materials).

3.1.4. Power calculations

Prior to conducting Study 2, we ran Studies S1 and S2 (see Supplemental materials) to compare the effects of ideal partner preference-matching on targets who are friends/acquaintances vs. current romantic partners, and for ideals for traits vs. behaviors. Studies S1 and S2 had relatively small sample sizes (for traits, $N = 110$ and $N = 109$ participants, respectively, and for behaviors, $N = 111$ participants in both studies), and self-generated ideals did not significantly predict romantic interest more strongly than other-generated ideals. However, those two studies revealed a nontrivial difference between self- vs. other-generated ideal *behaviors* only when participants were reporting on a current romantic partner (in terms of participants, $n = 52$ and $n = 95$ in Studies S1 and S2, respectively). In Study S1, the difference between self- and other-generated ideal behaviors for romantic partners was $\beta_{difference} = 0.11$ using our primary attribute-category exclusion approach, and in Study S2 this difference was $\beta_{difference} = 0.12$. These are approximately "small" effect sizes (Cohen, 1988), and despite the small sample sizes, we viewed this effect as an important avenue for additional testing. To conduct a power analysis, we used the data we ultimately obtained in Study S2 for the other-generated-behavior/romantic-interest association and the self-generated-behavior/other-generated-behavior association as inputs to solve for the sample size needed to provide 80% power to detect an effect of size $\beta_{difference} = 0.12$. Thus, assuming an other-generated-behavior/romantic-interest association of $\beta = 0.33$ and a self-generated-behavior/other-generated-behavior association of $\beta = 0.48$, $n = 448$ gives us 80% power to detect an effect of size $\beta_{difference} = 0.12$, which is a small effect (Cohen, 1992). Thus, we pursued a large sample in Study 2 with the intention of collecting enough participants reporting on a romantic partner that we would be powered to detect a difference of this size. We ended up achieving $n = 503$ in the romantic partners (behaviors) analysis, which provides 85% power to detect an effect of size $\beta_{difference} = 0.12$, and we ended up achieving $N = 595$ for the primary (behaviors) analysis.

3.2. Results

3.2.1. Primary analyses

The primary analyses for Study 2 were conducted on $N = 595$ participants (reporting on 2956 targets). Because participants were randomly assigned to a yoked-pair dyad and rated five separate targets, our data were structured such that each row represented a target with separate variables that kept track of (a) which participants constituted a yoked-pair, and (b) which five targets were nominated by which participant. The hierarchical nature of the data led us to test different multilevel models using the lme4 package (Bates et al., 2015) in the R environment (R Core Team, 2018). These included a three-level random intercepts, fixed slopes model that nested targets (level 1) within participant (level 2) within yoked-pair dyad (level 3), and a simpler two-level random intercept, fixed slopes model with targets (level 1) nested within participant (level 2) that ignored the yoked dyad. Given that the variances of the random effects were often small, we also tested a simple linear regression model. The two-level model specification nesting targets within participants provided the best fit to our data in terms of the lowest BIC values, and thus we present results from this model in our primary analyses (see Supplemental materials for summary of results across models tested for primary analyses, and see link to "JESP Models Tested Studies 1 and 2.html" document with R code and results from all models tested: https://osf.io/7fq9n/?view_only=336e96f6730b480182be7afe49383d31). As in Study 1, we standardized all variables, and then ran two separate models: (a) a self-generated model that predicted participants' romantic interest in their target from the extent to which their self-generated ideal attributes were rated as characteristic of their target, and (b) an other-generated model that predicted participants' romantic interest in their target from the extent to which their other-generated ideal attributes were rated as characteristic of their target. Data used in the primary analyses are available at https://osf.io/k28vb/?view_only=21cec00531ef4a19a56d2f26eae9b61a.

3.2.1.1. Traits. Consistent with the results from Study 1, both self-generated ideal trait ratings, $\beta = 0.39$, $t(2617.0) = 22.04$, $p < .001$, and other-generated ideal trait ratings, $\beta = 0.38$, $t(2306.0) = 21.33$, $p < .001$, positively predicted romantic interest (see Table 3 for descriptive statistics of all variables used in models). Once again, however, the Lee and Preacher (2013) test revealed no significant difference between the strength of these two associations, $z = 0.30$, $p = .761$ (see Fig. 1 and Table 4 for all associations and *N*s required for the Lee & Preacher, 2013, tests). Furthermore, sex did not moderate the associations between self-generated ideal trait ratings and romantic interest, $\beta = -0.05$, $t(2539.4) = 1.41$, $p = .160$, nor between other-generated ideal trait ratings and romantic interest, $\beta = -0.04$, $t(2060.1) = 0.92$, $p = .356$.

Because the Lee and Preacher (2013) test may not generalize to Study 2's multilevel context, we also used multilevel structural equation modeling to compare the self-generated ideal associations to the other-generated ideal associations while controlling for the correlation between the self and other ideal associations (Rosseel, 2017). Using R's lavaan package (Rosseel, 2012), we first fit an unconstrained model in which self-generated ideal trait ratings ($\beta = 0.50$, $SE = 0.15$, $p = .001$) and other-generated ideal trait ratings ($\beta = 0.60$, $SE = 0.17$, $p < .001$) simultaneously predicted romantic interest. All variables in the regression were latent and measured by manifest indicators at the within-participant level (i.e., three self-generated ideal trait ratings, three other-generated ideal trait ratings, and six items capturing romantic interest in a target). We then fit a constrained version of the model that forced the self-generated ideal association to be equal to the other-generated ideal association (β's $= 0.55$, $SEs = 0.03$, $p < .001$). If self-

Table 3
Descriptive statistics of primary variables: Study 2.

Variable	Mean	SD
Romantic interest (dependent variable)	4.52	3.19
Self-generated ideal trait ratings (independent variable for self-generated trait model)	7.87	2.25
Other-generated ideal trait ratings (independent variable for other-generated trait model)	7.75	2.20
Self-generated ideal behavior ratings (independent variable for self-generated behavior model)	7.65	2.38
Other-generated ideal behavior ratings (independent variable for other-generated behavior model)	7.46	2.43

Note. Means and standard deviations from the unstandardized variables used in the self-generated multilevel model and the other-generated multilevel model. All variables were measured on a scale from 1 (*strongly disagree)* to 11 (*strongly agree*).

Table 4
Summary of associations using the primary, attribute-category exclusion approach: Study 2.

Analysis	Self-generated-attribute/romantic-interest association	Other-generated-attribute/romantic-interest association	Self-generated-attribute/other-generated-attribute association	Sample Size
Primary (traits)	0.39	0.38	0.61	595
Primary (behaviors)	0.38	0.38	0.62	595
Friends/acquaintances (traits)	0.30	0.28	0.58	591
Friends/acquaintances (behaviors)	0.31	0.29	0.57	591
Romantic partners (traits)	0.53	0.45	0.60	501
Romantic partners (behaviors)	0.51	0.49	0.59	503
Single participants (traits)	0.36	0.36	0.57	155
Single participants (behaviors)	0.40	0.37	0.59	155
Committed participants (traits)	0.38	0.37	0.62	440
Committed participants (behaviors)	0.38	0.37	0.62	440

Note. The key associations (i.e., multilevel regression βs) used for the Lee and Preacher (2013) tests across our primary attribute-category exclusion analyses in Study 2. Self-generated- attribute/romantic-interest association refers to the association between self-generated ideal attribute ratings and romantic interest. Other-generated-attribute/romantic-interest association refers to the association between other-generated ideal attribute ratings and romantic interest. Self-generated-attribute/Other-generated-attribute association refers to the association between self-generated ideal attribute ratings and other-generated ideal attribute ratings. Sample size refers to the N who contributed to at least one of the three associations required for the Lee and Preacher (2013) test.

generated ideal trait ratings predict romantic interest more strongly than other-generated ideal trait ratings (our key hypothesis), then the unconstrained model should provide a better fit to our data than the constrained model. However, our results indicated that the unconstrained model did not fit the data any better than the constrained model, $\chi^2(1) = 0.09$, $p = .766$. Following Wagenmakers (2007), we then computed a Bayes Factor for this analysis by comparing the BIC values of the unconstrained and constrained models. With equal priors on the unconstrained and constrained models, the Bayes factor for this analysis was 39.4 (i.e., the posterior probability of the null hypothesis was approximately 0.98). According to Wagenmakers (2007), this value constitutes “strong” evidence for the null hypotheses (see legend in Fig. 1 and Supplemental materials). Thus, this analysis allows us to conclude that we have strong evidence to support the idea that self-generated ideal trait ratings do not predict romantic interest more strongly than other-generated ideal trait ratings. (Applying this analysis to the Study 1 data revealed a Bayes factor = 7.5, which constitutes "positive" evidence for the null hypothesis; see Fig. 1.)

3.2.1.2. Behaviors. We conducted the same analyses with our behavior data. Both self-generated ideal behavior ratings, $\beta = 0.38$, $t(2546.0) = 22.07$, $p < .001$, and other-generated ideal behavior ratings, $\beta = 0.38$, $t(2377.0) = 21.54$, $p < .001$, positively predicted romantic interest, and there was again no significant difference between the strength of these two associations, $z = 0.18$, $p = .854$ (see Fig. 1). Furthermore, sex did not moderate the associations between self-generated ideal behavior ratings and romantic interest, $\beta = -0.02$, $t(2340.3) = 0.56$, $p = .574$, nor between other-generated ideal behavior ratings and romantic interest, $\beta = 0.02$, $t(1951.8) = 0.64$, $p = .523$. Using multilevel structural equation modeling, our results again indicated that the unconstrained model in which the self- and other-generated ideal associations with romantic interest were allowed to differ did not fit the data any better than the constrained model in which these associations were assumed to be the same, $\chi^2(1) = 0.79$, $p = .376$. The Bayes factor for behaviors was 34.7 (i.e., the posterior probability of the null hypothesis was approximately 0.97), which constitutes “strong” evidence for the null hypotheses. Thus, we have strong evidence to support the idea that self-generated ideal behavior ratings do not predict romantic interest more strongly than other-generated ideal behavior ratings.

Participants expressed more romantic interest in their targets to the extent that they perceived them to possess positive traits and behaviors; however, it did not matter whether those traits or behaviors matched the ideals participants desired most in a partner or the ideals that another random participant desired most in a partner. Thus, as in Study 1, there was no evidence for the unique role of ideal partner preference-matching in predicting romantic interest.

3.2.2. Relationship status analyses

We examined whether predictive validity varied depending on (a) whether participants described the *target* as a romantic partner or not, and (b) whether participants described *themselves* as single or in a committed relationship. For targets who were not in a romantic relationship with the participant or a relationship was not reported (target $n = 2455$, participant $n = 591$), we used the same two-level random intercept, fixed slopes model with targets (level 1) nested within participant (level 2), as described above. For targets who were involved in a romantic relationship with the participant (target $n = 503$, participant $n = 437$), we used simple linear regressions because participants tended to report only one romantic partner (i.e., the average participant reported on $M = 1.15$ romantic partners). This value is larger than 1.00 because this subsample included (a) some single participants who had at least one romantic partner (e.g., a boyfriend/girlfriend; 8% of the subsample), and (b) committed participants who were in a romantic relationship with more than one target (10%).

We also examined whether predictive validity varied across all targets for single participants (target $n = 773$, participant $n = 155$) versus committed participants (target $n = 2183$, participant $n = 440$) as defined by the relationship status variable that characterizes participants themselves. This analysis examined $M = 4.96$–4.99 targets per participant, and so we used the same two-level random intercept, fixed slopes model with targets (level 1) nested within participant (level 2), as described above.

3.2.2.1. Friends and acquaintances (traits). Both self-generated ideal trait ratings, $\beta = 0.30$, $t(2402.5) = 16.92$, $p < .001$, and other-generated ideal trait ratings, $\beta = 0.28$, $t(2439.8) = 15.24$, $p < .001$, positively predicted romantic interest. Once again, however, the Lee and Preacher (2013) test revealed no significant difference between the strength of these two associations, $z = 0.56$, $p = .577$; see Fig. 1. Using multilevel structural equation modeling, our results again indicated that the unconstrained model did not fit the data any better than the constrained model, $\chi^2(1) = 1.0$, $p = .318$. The Bayes factor for traits was 22.7 (i.e., the posterior probability of the null hypothesis was approximately 0.96), which constitutes “strong” evidence for the null hypotheses.

3.2.2.2. Friends and acquaintances (behaviors). Both self-generated ideal behavior ratings, $\beta = 0.31$, $t(2426.4) = 17.16$, $p < .001$, and other-generated ideal behavior ratings, $\beta = 0.29$, $t(2438.2) = 15.54$, $p < .001$, positively predicted romantic interest. Again, the Lee and Preacher (2013) test revealed no significant difference between the strength of these two associations, $z = 0.55$, $p = .579$; see Fig. 1. Using multilevel structural equation modeling, our results indicated that the constrained model actually provided a better fit to the data than the unconstrained, $\chi^2(1) = 3.96$, $p = .047$, which is even stronger evidence that the self-generated and other-generated paths do not differ in this analysis. In addition, the Bayes factor for behaviors was 6.4 (i.e., the posterior probability of the null hypothesis was approximately 0.87), which constitutes "positive" evidence for the null hypotheses.

3.2.2.3. Romantic partners (traits). Both self-generated ideal trait ratings, $\beta = 0.53$, $SE = 0.04$, $p < .001$, and other-generated ideal trait ratings, $\beta = 0.45$, $SE = 0.04$, $p < .001$, positively predicted romantic interest. In this sample, the Lee and Preacher (2013) test revealed that self-generated ideal trait ratings predicted romantic interest more strongly than other-generated ideal trait ratings, $z = 2.37$, $p = .018$; see Fig. 1 (see Supplemental materials for an additional robustness check of this significant result.) However, using structural equation modeling, our results indicated that the unconstrained model did not fit the data any better than the constrained model, $\chi^2(1) = 0.18$, $p = .675$. The Bayes factor for traits was 15.7 (i.e., the posterior probability of the null hypothesis was approximately 0.94), which constitutes "positive" evidence for the null hypotheses.

3.2.2.4. Romantic partners (behaviors). Both self-generated ideal behavior ratings, $\beta = 0.51$, $SE = 0.04$, $p < .001$, and other-generated ideal behavior ratings, $\beta = 0.49$, $SE = 0.04$, $p < .001$, positively predicted romantic interest. Once again, however, the Lee and Preacher (2013) test revealed no significant difference between the strength of these two associations, $z = 0.59$, $p = .554$; see Fig. 1. Using structural equation modeling, our results indicated that the unconstrained model did not fit the data any better than the constrained model, $\chi^2(1) = 0.07$, $p = .785$. The Bayes factor for traits was 20.7 (i.e., the posterior probability of the null hypothesis was approximately 0.95), which constitutes "strong" evidence for the null hypotheses.

3.2.2.5. Single participants (traits). For single participants, both self-generated ideal trait ratings, $\beta = 0.36$, $t(770.8) = 10.83$, $p < .001$, and other-generated ideal trait ratings, $\beta = 0.36$, $t(749.8) = 10.19$, $p < .001$, positively predicted romantic interest. Once again, however, the Lee and Preacher (2013) test revealed no significant difference between the strength of these two associations, $z = 0.04$, $p = .965$; see Fig. 1. Using multilevel structural equation modeling, our results again indicated that the unconstrained model did not fit the data any better than the constrained model, $\chi^2(1) = 1.50$, $p = .221$. The Bayes factor for traits was 9.8 (i.e., the posterior probability of the null hypothesis was approximately 0.91), which constitutes "positive" evidence for the null hypotheses.

3.2.2.6. Single participants (behaviors). For single participants, both self-generated ideal behavior ratings, $\beta = 0.40$, $t(769.9) = 11.91$, $p < .001$, and other-generated ideal behavior ratings, $\beta = 0.37$, $t(766.8) = 10.72$, $p < .001$, positively predicted romantic interest. But again, the Lee and Preacher (2013) test revealed no significant difference between the strength of these two associations, $z = 0.45$, $p = .652$; see Fig. 1. Using multilevel structural equation modeling, our results again indicated that the unconstrained model did not fit the data any better than the constrained model, $\chi^2(1) = 2.51$, $p = .113$. The Bayes factor for behaviors was 7.6 (i.e., the posterior probability of the null hypothesis was approximately 0.88), which constitutes "positive" evidence for the null hypotheses.

3.2.2.7. Committed participants (traits). For committed participants, both self-generated ideal trait ratings, $\beta = 0.38$, $t(2183.0) = 19.08$, $p < .001$, and other-generated ideal trait ratings, $\beta = 0.37$, $t(2179.0) = 18.55$, $p < .001$, positively predicted romantic interest. Once again, however, the Lee and Preacher (2013) test revealed no significant difference between the strength of these two associations, $z = 0.26$, $p = .793$; see Fig. 1. Using multilevel structural equation modeling, our results again indicated that the unconstrained model did not fit the data any better than the constrained model, $\chi^2(1) = 0.79$, $p = .373$. The Bayes factor for traits was 23.9 (i.e., the posterior probability of the null hypothesis was approximately 0.96), which constitutes "strong" evidence for the null hypotheses.

3.2.2.8. Committed participants (behaviors). For committed participants, both self-generated ideal behavior ratings, $\beta = 0.38$, $t(2180.0) = 18.88$, $p < .001$, and other-generated ideal behavior ratings, $\beta = 0.37$, $t(2178.0) = 18.63$, $p < .001$, positively predicted romantic interest. But again, the Lee and Preacher (2013) test revealed no significant difference between the strength of these two associations, $z = 0.26$, $p = .793$; see Fig. 1. Using multilevel structural equation modeling, our results again indicated that the unconstrained model did not fit the data any better than the constrained model, $\chi^2(1) = 0.06$, $p = .811$. The Bayes factor for traits was about 42.7 (i.e., the posterior probability of the null hypothesis was approximately 0.98), which constitutes "strong" evidence for the null hypotheses.

3.3. Discussion

The overall results of Study 2 do not support the hypothesis that self-generated ideals predict romantic interest more strongly than other-generated ideals. In the primary analyses, the friends and acquaintances analyses, and the single and committed participant analyses, we found no evidence of ideal partner preference-matching for either traits or behaviors (Fig. 1, data points 2–5 and 8–11). In the subsample of romantic partners—which guided our power analysis—we did find that self-generated ideal traits were more strongly associated with romantic interest than other-generated ideal traits, but this effect size was very small (Fig. 1, data point 6, $\beta_{difference} = 0.08$), and analyses with structural equation modeling suggested that this effect was not robust. Moreover, the null findings for ideal behaviors in this same subsample (Fig. 1, data point 7, $\beta_{difference} = 0.02$), and the fact that this significant effect fell to marginal or nonsignificance using all of the alternative exclusion approaches (see Supplemental materials) suggests the possibility that the one significant effect in Fig. 1 may be a Type I error.

By using BIC values and Structural Equation Modeling, we were able to quantify the strength of the evidence for the null hypothesis in this study. Out of the 30 times both the constrained and unconstrained models converged across our attribute-category exclusion analyses, our synonym-level exclusion analyses, our Fletcher three-factor exclusion analyses, and analyses using all data (no exclusions for duplicate attributes), 50.0% of the time we found "strong" evidence for the null hypothesis, 46.7% of the time we found "positive" evidence for the null hypothesis, and 3.3% of the time we found no evidence for the null hypothesis (see summary of Bayes Factors for Study 2 in Figs. 1, S3, S4, and S5). In other words, we have "strong" or "positive" evidence for the null hypothesis in 96.7% of our Study 2 analyses.

4. General discussion

The current studies offer focused tests of the hypothesis that ideal partner preference-matching predicts romantic interest. Specifically, the present studies used a novel experimental approach with a yoked control to isolate the unique predictive power of idiosyncratic participant-generated ideals from the normative desirability of those ideals. Analyses generally failed to reveal evidence for this hypothesis, either for new potential partners (Study 1) or existing acquaintances, friends, and romantic partners (Study 2): Participants expressed more romantic interest in targets to the extent that they thought those targets had positive attributes, but there was no boost in predictive power if those attributes represented one's own ideals or the ideals of a random other participant.

Our various exclusion schemes maximized idiosyncratic variability in self- versus other-generated ideals at different levels of breadth vs. precision (see Appendix A and Table 1). Using a precise, well-powered attribute-category exclusion approach in Study 2 (with our largest sample sizes), each estimated difference between the predictive power of self- vs. other-generated ideal attribute ratings was indeed positive (Fig. 1); that is, self-generated ideals predicted romantic interest more strongly than other-generated ideals. Yet, despite this directional trend, just one significant difference reliably emerged in our analyses, and it survived no additional robustness checks (e.g., SEM analysis). Thus, although the overall effect size might indeed be larger than zero, our results suggest that it is extremely small (e.g., the average difference between self- vs. other-generated ideals across Studies 1 and 2 is $\beta_{\text{difference}} = 0.03$; see Table 1). In analyses using the synonym-level approach and the Fletcher three-factor exclusion approach, we again found no reliable evidence that self-generated ideal attributes predict romantic interest more strongly than other-generated ideal attributes, although these approaches remove more participants (see Table 1) and are therefore less well powered than the primary approach. In analyses using all data (no exclusions for duplicate attributes; see Supplemental materials), we also found no reliable evidence that self-generated ideal attributes predict romantic interest more strongly than other-generated ideal attributes. Whether a given scholar a priori favors a highly powered narrow exclusion strategy (i.e., attribute-level), a moderately powered broad exclusion strategy (i.e., factor-level), or something in between (i.e., the synonym-level exclusion strategy), the average self- vs. other-generated effect size is essentially identical across analyses ($\beta_{\text{difference}} = 0.03$–$0.04$). The experimental approach developed in the present studies offers a new technique for testing matching hypotheses, and our results suggest that if one wants to claim evidence for an effect, it may be necessary to use a sample size well beyond what is typically used in social psychology.

4.1. Implications

4.1.1. Adding clarity to the ideal partner preference-matching literature

Interpretive confusions pervade the literature on ideal partner preference-matching, so it is essential to clarify how the current studies fit into this broader literature. For example, some recent studies have begun to interpret the simple correlation between a person's ideals and a current partner's traits as evidence for the predictive validity of ideal partner preference-matching (e.g., Campbell, Chin, & Stanton, 2016; Conroy-Beam & Buss, 2016; Gerlach, Arslan, Schultze, Reinhard, & Penke, 2019), but this procedure does not test the predictive validity of the “match” at all. Indeed, there are six ways that ideal-partner trait correlations can emerge even if ideal-matching has no effect on how positively people evaluate a partner (Eastwick, Finkel et al., 2019). For example, people draw from the attributes of potential partners who are generally in their social milieu when they form their ideals (Eastwick, Smith, & Ledgerwood, 2019); also, ideals are correlated with perceiver effects, which are general tendencies to believe that others possess particular attributes, (e.g., Rau et al., 2019). The results of the current studies suggest that ideal-partner trait correlations are more likely to be caused by one or more of these alternative mechanisms than by a process whereby participants positively evaluate potential partners who match (rather than mismatch) their ideals.

Colloquially speaking, the current studies examined whether individual differences in the attributes that participants say they prefer in a partner uniquely predict their interest in partners who possess those attributes. After addressing the normative desirability of positive traits and behaviors (via the yoked-control comparison), what participants said they preferred did not have any special correspondence with their romantic interest ratings. This finding raises questions about the way that people use their self-reported ideals as a comparison standard when evaluating partners, and the data are consistent with the possibility that people exercise considerable flexibility in the way they interpret a partner's traits after a live interaction has taken place (Eastwick et al., 2011; Sunnafrank & Miller, 1981). But importantly, the dependent variable in these studies was a romantic evaluation of a blind date partner or partner/friend/acquaintance, not the choice to become romantic partners with one person instead of another (indeed, to our knowledge, no studies of real-life romantic evaluations have used such a choice measure).

More formally, the current research question illustrates the distinction between *summarized preferences* for attributes (e.g., the attributes people say they like in the abstract; Anderson, 1965) and *functional preferences* for attributes (e.g., the attributes that drive people's liking across a set of targets; Lawless & Heymann, 2010; Wood & Brumbaugh, 2009). That is, ideal partner preferences are illustrations of summarized preferences, and the actual associations between attribute ratings and romantic interest are illustrations of functional preferences (Ledgerwood, Eastwick & Smith, 2018). These two preferences occupy different levels of abstraction, and they may have different consequences. Across the domains in which summarized and functional preferences have been studied (e.g., food and beverage preferences, personnel selection), they seem to correspond more highly when the targets are simple rather than complex (see Ledgerwood, Eastwick, & Smith, 2018, for a review). The complexity inherent in the process of romantically evaluating other real-life partners may explain why correspondence between the attributes that people say they like and the attributes they actually like seems to be particularly poor in the studies we conducted here.

4.1.2. Individual differences and normative ideal partner preferences

The current studies were designed to address individual differences in ideal partner preferences: That is, what does it mean that Faye's ideal consists of attributes like kind, intelligent, and attractive, whereas Sonia's ideal consists of attributes like ambitious, generous, and adventurous? Indeed, idiosyncratic differences are the basis of nearly all studies of ideals, including studies examining associations between ideals and self-rated attributes (e.g., Figueredo, Sefcek, & Jones, 2006), ratings of others (e.g., Gerlach et al., 2019), and other mating-relevant individual differences (e.g., sociosexuality; Simpson & Gangestad, 1992). Idiosyncratic differences are similarly the basis of classic studies that highlight sex as a predictor of mate preferences (e.g., Buss, 1989); that is, biological sex is a categorical variable that predicts that some people rate an attribute higher than other people, just like a continuous variable would. If individual differences in ideal partner preferences have weak or near-zero unique predictive effects on romantic evaluations, as the current findings tentatively suggest, then these oft-studied individual differences likely affect mate evaluation and choice via mechanisms other than mate preferences (Eastwick et al., 2014).

However, the current studies were not designed to address normative ideal partner preferences: What does it mean that a population of participants rates trustworthiness higher than earning potential in an ideal partner (Fletcher et al., 1999)? In principle, it is possible to examine the predictive validity of both normative and idiosyncratic components of ideals simultaneously. For example, in the domain of workplace preferences for attributes like *flexibility* and *low level of conflict*, Wood et al. (2019) found that people are indeed happier with their jobs when their jobs match their ideal preferences. But this finding is driven not by idiosyncratic preference-matching but rather by normative preference-matching; people like jobs that have features of the normatively ideal job. At a conceptual level, Wood et al.'s (2019) findings highlight the importance of subtracting the normative desirability confound if researchers intend to argue for the incremental value of assessing and modeling individual differences in self-reported preferences. Thus, prior studies that failed to account for this issue have ambiguous relevance for the predictive validity of individual differences in preferences (Conroy-Beam et al., 2016; Fletcher et al., 1999, Study 6; Fletcher et al., 2000)—including papers by some authors of the current manuscript (e.g., Eastwick et al., 2011, Study 3).

4.2. Future directions

It is perhaps noteworthy that the largest effect (albeit still small, $\beta_{\text{difference}} = 0.08$) tended to emerge when rating ideal traits for a current romantic partner (see Fig. 1). This is consistent with findings that the most reliable effects of ideal partner preference-matching in the prior literature tended to examine people's ongoing romantic relationships (e.g., Fletcher et al., 2000, 1999). However, the present studies (which removed normative-desirability confounds) suggest that this effect may be smaller and less reliable than implied by the published literature. Indeed, structural equation modeling analyses suggested that this $\beta_{\text{difference}} = 0.08$ effect for current romantic partners (traits) may itself be inflated or non-systematic. Future work could focus on this subsample and recruit a much larger sample size to provide an even stronger test of ideal partner preference-matching.

Ideals for behaviors in Study 2 also did not reveal evidence of predictive validity; that is, the behavior rating association with romantic interest was effectively identical for self and other-nominated ideal behaviors. This null result was somewhat surprising to us in light of prior research and theory suggesting that, over time, people develop particular expectations about the way their romantic partners behave (Lakey & Orehek, 2011; McNulty & Karney, 2004). However, it is also possible that ideals matter on a moment-to-moment basis within the context of the particular relationship in which those standards formed. That is, I might ideally want my current partner to have breakfast with me every day—and I would indeed be disappointed on days when he/she does not engage in this behavior—but this ideal would not apply to any past or alternative relationship. Future research might consider whether ideal-matching—whether for behaviors or even for traits—exerts effects on a moment-to-moment basis when those ideals are bound to the context of a specific partnership.

One limitation of the present research is that it is constrained to the traits and behaviors that people bring to mind spontaneously when asked about ideals. Ideal partner preference-matching effects could be strong for attributes that our participants never nominated (e.g., “someone of my own social class,” “someone within 15 years of my own age”); that is, the average participant might not be fully aware of the ideals that idiosyncratically matter to him or her (Wilson, Laser, & Stone, 1982). Future studies using other techniques could capture the extent to which ideals influence romantic partner choice without requiring participants to articulate exactly what those ideals are (e.g., by examining the degree of idiosyncratic consistency in romantic partner choices over time; Eastwick, Harden, Shukusky, Morgan, & Joel, 2017). On the other hand, it is also possible that people are fully aware of the ideals that matter to them and can articulate them clearly, but that another limitation of our design is that it artificially constrained people to nominate an incomplete subset (i.e., their top three ideals) out of the larger set of ideals that truly drives their romantic interest. Although prior studies relying on large numbers of researcher-generated attributes have tended to provide ambiguous results (which, in turn, inspired us to come up with this new yoked control design), it might be useful for future research to implement a hybrid approach that asks participants to self-nominate a larger number of ideals. Finally, future studies might also find support for predictive validity by combining (a) a nomination procedure and yoked design (as in the current studies) with (b) a procedure that encourages participants to nominate attributes that they value but are considerably more idiosyncratic (e.g., traits like *proud* and *satirical*, as used in Eastwick et al., 2011, Study 2).

4.3. Conclusion

This programmatic set of studies implemented a focused test of the unique influence of ideal partner preference-matching on romantic interest, and generally failed to observe such influences. To be sure, the ideals that participants nominated tended to be attributes that predicted romantic interest positively—and often quite strongly—but their self-nominated ideals tended not to have any additional influence beyond the positivity of the attributes themselves. Overall, these findings suggest that individual differences in the attributes that people say they desire in a romantic partner may provide little insight into the attributes that uniquely inspire their romantic interest in real-life partners and potential partners.

Open practices

The Study 2 data used for the primary models can be found at the following link: https://osf.io/k28vb/?view_only=21cec00531ef4a19a56d-2f26eae9b61a. We cannot post the data for Study 1, unfortunately, because the dyadic nature of the data means that they are identifiable to participants themselves (e.g., participants could look up what their blind date partners said about them).

Appendix A

	Attribute-Category	Overall	Study 1	Study 2	Grouping for exclusion scheme 2 (synonym)	Grouping for exclusion scheme 3 (three-factor)
1	Good Sense of Humor	12.57%	11.25%	12.99%	none	Vitality/Attractiveness
2	Intelligent	9.54%	7.36%	10.24%	1	none
3	Honest	7.19%	4.09%	8.11%	2	Warmth/Trustworthiness
4	Attractive	6.54%	13.09%	5.15%	3	Vitality/Attractiveness
5	Kind	5.15%	3.27%	5.79%	4	Warmth/Trustworthiness
6	Understanding	4.47%	1.23%	5.36%	4	Warmth/Trustworthiness
7	Ambitious	3.82%	3.48%	3.47%	none	Vitality/Attractiveness
8	Loyalty	3.54%	0.82%	3.99%	2	Warmth/Trustworthiness
9	Caring	3.20%	1.23%	3.26%	4	Warmth/Trustworthiness
10	In Love (feelings)	2.52%	2.66%	2.75%	5	Warmth/Trustworthiness
11	Trustworthy	2.27%	2.25%	1.93%	2	Warmth/Trustworthiness
12	Considerate	1.56%	1.02%	1.54%	4	Warmth/Trustworthiness
13	Good Fun	1.56%	1.84%	1.54%	none	Vitality/Attractiveness
14	Reliable	1.39%	1.64%	1.24%	2	Warmth/Trustworthiness
15	Patient	1.30%	0.61%	1.46%	none	none
16	Warm	1.30%	1.43%	1.24%	4	Warmth/Trustworthiness
17	Outgoing	1.19%	2.04%	0.94%	none	Vitality/Attractiveness
18	Stable	1.16%	1.23%	1.33%	2	Warmth/Trustworthiness
19	Confident	1.13%	1.02%	0.99%	9	Vitality/Attractiveness
20	Nice Body	1.13%	1.23%	0.99%	3	Vitality/Attractiveness
21	Adventurous	1.08%	1.23%	0.90%	5	Vitality/Attractiveness

22	Generous	1.08%	0.20%	1.33%	4	none
23	Passionate	1.05%	1.64%	1.07%	5	Vitality/Attractiveness
24	Broad-Minded	1.01%	1.84%	0.89%	none	Warmth/Trustworthiness
25	Religious Beliefs	0.99%	0.82%	0.64%	none	none
26	Compatibility (thinking/talking/beliefs)	0.96%	5.93%	0.21%	none	none
27	Easygoing	0.93%	0.41%	1.16%	6	Warmth/Trustworthiness
28	Personality/Similar Personalities	0.91%	3.48%	0.39%	none	Vitality/Attractiveness
29	Respect	0.91%	0.20%	0.82%	4	none
30	Sporty and Athletic	0.88%	0.20%	1.03%	7	Vitality/Attractiveness
31	Creative	0.85%	0.82%	0.90%	1	none
32	Sensitive	0.85%	0.00%	0.99%	8	Warmth/Trustworthiness
33	Friendly	0.79%	1.84%	0.60%	4	Warmth/Trustworthiness
34	Interesting	0.76%	2.25%	0.51%	5	Vitality/Attractiveness
35	Supportive	0.74%	0.61%	0.77%	none	Warmth/Trustworthiness
36	Affectionate	0.68%	0.41%	0.69%	4	Warmth/Trustworthiness
37	Independent	0.68%	0.82%	0.60%	none	Vitality/Attractiveness
38	Sexy	0.65%	0.00%	0.90%	3	Vitality/Attractiveness
39	Romantic	0.62%	0.20%	0.69%	5	Warmth/Trustworthiness
40	Similar Interests	0.62%	1.02%	0.51%	none	none
41	Financially Secure	0.57%	0.00%	0.64%	none	Status/Resources
42	Humble	0.54%	1.43%	0.43%	4	none
43	Relaxed	0.54%	0.61%	0.51%	6	Vitality/Attractiveness
44	Educated	0.51%	0.41%	0.47%	1	none
45	Hopeful/Optimistic	0.51%	0.82%	0.51%	9	none
46	Family-/Friend-Oriented	0.48%	0.41%	0.43%	none	none
47	Active Lifestyle	0.45%	0.61%	0.39%	7	Vitality/Attractiveness
48	Communicative	0.45%	1.02%	0.43%	none	Warmth/Trustworthiness
49	Moral	0.42%	0.61%	0.47%	none	none
50	Refined	0.40%	0.20%	0.56%	none	none
51	Good Lover	0.37%	0.00%	0.47%	none	Vitality/Attractiveness
52	Masculinity	0.34%	0.00%	0.30%	none	none
53	Spontaneous	0.34%	0.82%	0.30%	6	Vitality/Attractiveness
54	Assertive	0.31%	0.00%	0.30%	9	Vitality/Attractiveness
55	Good Listener	0.31%	0.20%	0.26%	none	Warmth/Trustworthiness
56	Self-Aware	0.31%	1.23%	0.17%	none	Warmth/Trustworthiness
57	Acceptance	0.28%	0.00%	0.39%	none	Warmth/Trustworthiness
58	Successful	0.28%	0.00%	0.13%	none	Status/Resources
59	Likes Children	0.23%	0.41%	0.21%	none	none
60	Commitment	0.20%	0.41%	0.17%	none	Warmth/Trustworthiness
61	Confronts Conflict	0.17%	0.20%	0.13%	none	none
62	Geeky	0.17%	0.41%	0.17%	none	none
63	Mature	0.17%	0.41%	0.17%	none	Warmth/Trustworthiness
64	Nice House or Apartment/Rich	0.17%	0.00%	0.21%	none	Status/Resources
65	Dresses Well	0.14%	0.20%	0.09%	none	Status/Resources
66	Politics	0.14%	0.00%	0.21%	none	none
67	Protective	0.14%	0.00%	0.09%	4	none
68	Reserved	0.14%	0.00%	0.17%	none	none
69	Compromise	0.08%	0.00%	0.13%	none	none
70	Display Emotion	0.08%	0.00%	0.13%	8	none
71	Equality	0.08%	0.00%	0.04%	none	Warmth/Trustworthiness
72	Exciting	0.08%	0.41%	0.04%	5	Vitality/Attractiveness
73	Good Job	0.08%	0.00%	0.13%	none	Status/Resources
74	Unusual	0.08%	0.00%	0.13%	none	none
75	Animal Lover	0.06%	0.00%	0.09%	none	none
76	Appropriate Ethnicity	0.06%	0.41%	0.00%	none	Status/Resources
77	Challenging	0.06%	0.41%	0.00%	none	Vitality/Attractiveness
78	Goofy	0.06%	0.20%	0.04%	none	none
79	Sharing	0.06%	0.00%	0.09%	none	Warmth/Trustworthiness
80	Asexual	0.03%	0.00%	0.04%	none	none
81	Chaste	0.03%	0.20%	0.00%	10	none
82	Childfree	0.03%	0.00%	0.04%	none	none
83	Deals Well with Criticism	0.03%	0.00%	0.04%	none	none
84	Does Not Smoke	0.03%	0.20%	0.00%	none	none
85	Early riser	0.03%	0.00%	0.04%	none	none
86	Good Memory	0.03%	0.20%	0.00%	none	none
87	Location	0.03%	0.20%	0.00%	none	none
88	Neat	0.03%	0.00%	0.04%	none	none
89	Non-materialistic	0.03%	0.00%	0.04%	none	none
90	Popular	0.03%	0.20%	0.00%	3	none
91	Simple	0.03%	0.00%	0.04%	none	none
92	Soul	0.03%	0.00%	0.00%	5	none
93	Tattoos	0.03%	0.00%	0.04%	none	none

94	Appropriate Age	0.00%	0.00%	0.00%	none	Status/Resources
95	Monogamous	0.00%	0.00%	0.00%	10	Warmth/Trustworthiness

Note. The percentage of total ideals nominated at pre-test by attribute-category grouping across all studies. Attribute groupings are sorted from descending frequency on the Overall column. That is, across Studies 1, 2, S1, and S2, the three most frequently nominated attributes are Good Sense of Humor, Intelligent, and Honest. Synonym-level exclusion groupings were devised using the most common (i.e., dark orange) synonyms and definitions from thesaurus.com (e.g., intelligent and creative are synonyms). Three-factor groupings were assigned by loadings over 0.40 in Fletcher et al. (1999) Table 1 (bolded values), and if the attribute-category was not found in Table 1, then we used the values reported in Fletcher et al. (1999) Table 2. The three-factor approach excludes ideals if they matched on one of the three factors from Fletcher et al. (1999); if the ideals did not fit into one of the three factors, they were excluded if they matched at the attribute-category level for one of the remaining 40 attributes (as in our primary, attribute-category exclusion approach).

Appendix B. Supplemental materials

Supplemental materials to this article can be found online at https://doi.org/10.1016/j.jesp.2020.103968.

References

Anderson, N. H. (1965). Averaging versus adding as a stimulus-combination rule in impression formation. *Journal of Experimental Psychology, 70*, 394–400.
Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*, 1–48.
Buss, D. M. (1989). Sex differences in human mate preferences: Evolutionary hypotheses tested in 37 cultures. *Behavioral and Brain Sciences, 12*, 1–49.
Campbell, L., Chin, K., & Stanton, S. C. (2016). Initial evidence that individuals form new relationships with partners that more closely match their ideal preferences. *Collabra: Psychology, 2*, 2.
Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.
Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155–159.
Conroy-Beam, D., & Buss, D. M. (2016). How are mate preferences linked with actual mate selection? Tests of mate preference integration algorithms using computer simulations and actual mating couples. *PLoS One, 11*, 1–22.
Conroy-Beam, D., & Buss, D. M. (2017). Euclidean distances discriminatively predict short-term and long-term attraction to potential mates. *Evolution and Human Behavior, 38*, 442–450.
Conroy-Beam, D., Goetz, C. D., & Buss, D. M. (2016). What predicts romantic relationship satisfaction and mate retention intensity: Mate preference fulfillment or mate value discrepancies? *Evolution and Human Behavior, 37*, 440–448.
Cronbach, L. J., & Gleser, G. C. (1957). *Psychological tests and personnel decisions.* Urbana, IL: University of Illinois Press.
DeBruine, L. M., Jones, B. C., Little, A. C., Boothroyd, L. G., Perrett, D. I., Penton-Voak, I. S., ... Tiddeman, B. P. (2006). Correlated preferences for facial masculinity and ideal or actual partner's masculinity. *Proceedings of the Royal Society of London B: Biological Sciences, 273*, 1355–1360.
Eastwick, P. W., & Finkel, E. J. (2008). Sex differences in mate preferences revisited: Do people know what they initially desire in a romantic partner? *Journal of Personality and Social Psychology, 94*, 245–264.
Eastwick, P. W., Finkel, E. J., & Eagly, A. H. (2011). When and why do ideal partner preferences affect the process of initiating and maintaining romantic relationships? *Journal of Personality and Social Psychology, 101*, 1012–1032.
Eastwick, P. W., Finkel, E. J., & Simpson, J. A. (2019). Best practices for testing the predictive validity of ideal partner preference-matching. *Personality and Social Psychology Bulletin, 45*, 167–181.
Eastwick, P. W., Harden, K. P., Shukusky, J. A., Morgan, T. A., & Joel, S. (2017). Consistency and inconsistency among romantic partners over time. *Journal of Personality and Social Psychology, 112*, 838–859.
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.
Eastwick, P. W., & Smith, L. K. (2018). Sex-differentiated effects of physical attractiveness on romantic desire: A highly powered, preregistered study in a photograph evaluation context. *Comprehensive Results in Social Psychology, 3*, 1–27.
Eastwick, P. W., Smith, L. K., & Ledgerwood, A. (2019). How do people translate their experiences into abstract attribute preferences? *Journal of Experimental Social Psychology, 85* 103837.
Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149–1160.
Figueredo, A. J., Sefcek, J. A., & Jones, D. N. (2006). The ideal romantic partner personality. *Personality and Individual Differences, 41*, 431–441.
Finkel, E. J., Simpson, J. A., & Eastwick, P. W. (2017). The psychology of close relationships: Fourteen core principles. *Annual Review of Psychology, 68*, 383–411.
Fletcher, G. J., Kerr, P. S., Li, N. P., & Valentine, K. A. (2014). Predicting romantic interest and decisions in the very early stages of mate selection: Standards, accuracy, and sex differences. *Personality and Social Psychology Bulletin, 40*, 540–550.
Fletcher, G. J. O., & Simpson, J. A. (2000). Ideal standards in close relationships: Their structure and functions. *Current Directions in Psychological Science, 9*, 102–105.
Fletcher, G. J. O., Simpson, J. A., & Thomas, G. (2000). Ideals, perceptions, and evaluations in early relationship development. *Journal of Personality and Social Psychology, 79*, 933–940.
Fletcher, G. J. O., Simpson, J. A., Thomas, G., & Giles, L. (1999). Ideals in intimate relationships. *Journal of Personality and Social Psychology, 76*, 72–89.
Fletcher, G. J., Tither, J. M., O'Loughlin, C., Friesen, M., & Overall, N. (2004). Warm and homely or cold and beautiful? Sex differences in trading off traits in mate selection. *Personality and Social Psychology Bulletin, 30*, 659–672.
Gable, S. L., Gosnell, C. L., Maisel, N. C., & Strachman, A. (2012). Safely testing the alarm: Close others' responses to personal positive events. *Journal of Personality and Social Psychology, 103*, 963–981.
Gerlach, T. M., Arslan, R. C., Schultze, T., Reinhard, S. K., & Penke, L. (2019). Predictive validity and adjustment of ideal partner preferences across the transition into romantic relationships. *Journal of Personality and Social Psychology, 116*, 313.
Gottman, J. M. (1998). Psychology and the study of marital processes. *Annual Review of Psychology, 49*, 169–197.
Joel, S., Teper, R., & MacDonald, G. (2014). People overestimate their willingness to reject potential romantic partners by overlooking their concern for other people. *Psychological Science, 25*, 2233–2240.
John, O. P., Hampson, S. E., & Goldberg, L. R. (1991). The basic level in personality-trait hierarchies: Studies of trait use and accessibility in different contexts. *Journal of Personality and Social Psychology, 60*, 348.
Jonason, P. K., Garcia, J. R., Webster, G. D., Li, N. P., & Fisher, H. E. (2015). Relationship dealbreakers: Traits people avoid in potential mates. *Personality and Social Psychology Bulletin, 41*, 1697–1711.
Kelley, H. H., Berscheid, E., Christensen, A., Harvey, J. H., Huston, T. L., Levinger, G., & Peterson, D. R. (1983). *Close relationships.* New York: Freeman265–314.
Kline, R. B. (2005). *Principles and practice of structural equation modeling.* New York, NY: Guilford Press.
Lakey, B., & Orehek, E. (2011). Relational regulation theory: A new approach to explain the link between perceived social support and mental health. *Psychological Review, 118*, 482–495.
Lam, B. C. P., Cross, S. E., Wu, T.-F., Yeh, K.-H., Wang, Y.-C., & Su, J. C. (2016). What do you want in a marriage? Examining marriage ideals in Taiwan and the United States. *Personality and Social Psychology Bulletin, 42*, 703–722.
Lawless, H. T., & Heymann, H. (2010). *Sensory evaluation of food: Principles and practices.* Springer Science & Business Media.
Ledgerwood, A., Eastwick, P. W., & Smith, L. K. (2018). Toward an integrative framework for studying human evaluation: Attitudes toward objects and attributes. *Personality and Social Psychology Review, 22*, 378–398.
Ledgerwood, A., & Shrout, P. E. (2011). The trade-off between accuracy and precision in latent variable models of mediation processes. *Journal of Personality and Social Psychology, 101*, 1174–1188.
Lee, I. A., & Preacher, K. J. (2013, September). Calculation for the test of the difference between two dependent correlations with one variable in common. [Computer software]. Available from http://quantpsy.org.
Li, N. P., Bailey, J. M., Kenrick, D. T., & Linsenmeier, J. A. (2002). The necessities and luxuries of mate preferences: Testing the tradeoffs. *Journal of Personality and Social Psychology, 82*, 947.
Li, N. P., & Kenrick, D. T. (2006). Sex similarities and differences in preferences for short-term mates: What, whether, and why. *Journal of Personality and Social Psychology, 90*, 468–489.
McNulty, J. K., & Karney, B. R. (2004). Positive expectations in the early years of marriage: Should couples expect the best or brace for the worst? *Journal of Personality and Social Psychology, 86*, 729–743.
McNulty, J. K., & Russell, V. M. (2010). When "negative" behaviors are positive: A contextual analysis of the long-term effects of problem-solving behaviors on changes in relationship satisfaction. *Journal of Personality and Social Psychology, 98*, 587.
Overall, N. C., Fletcher, G. J., & Simpson, J. A. (2006). Regulation processes in intimate relationships: The role of ideal standards. *Journal of Personality and Social Psychology, 91*, 662.
Przybylinski, E., & Andersen, S. M. (2015). Systems of meaning and transference: Implicit significant-other activation evokes shared reality. *Journal of Personality and Social Psychology, 109*, 636–661.
R Core Team (2018). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.
Rau, R., Carlson, E. N, Back, M. D., Barranti, M., Gebauer, J. E., Human, L. J, Leising, D., & Nestler, S. (2019). What is the structure of perceiver effects? On the importance of global positivity and trait-specificity across personality domains and judgment contexts. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/pspp0000278 Advance online publication.

Rogers, K. H., Wood, D., & Furr, R. M. (2018). Assessment of similarity and self-other agreement in dyadic relationships: A guide to best practices. *Journal of Social and Personal Relationships, 35*, 112–134.

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*, 1–36.

Rosseel, Y. (2017). Multilevel structural equation modeling with lavaan. Retrieved from https://users.ugent.be/~yrosseel/lavaan/zurich2017/MULTILEVEL/lavaan_multilevel_zurich2017.pdf.

Simpson, J. A., Fletcher, G. J. O., & Campbell, L. (2001). The structure and function of ideal standards in close relationships. In G. J. O. Fletcher, & M. Clark (Eds.). *The Blackwell handbook in social psychology: Interpersonal processes* (pp. 86–106). Oxford, England: Blackwell.

Simpson, J. A., & Gangestad, S. W. (1992). Sociosexuality and romantic partner choice. *Journal of Personality, 60*, 31–51.

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin, 87*, 245–251.

Sunnafrank, M. J., & Miller, G. R. (1981). The role of initial conversations in determining attraction to similar and dissimilar strangers. *Human Communication Research, 8*, 16–25.

Thibaut, J. W., & Kelley, H. H. (1959). *The social psychology of groups.* New York: Wiley.

Wagenmakers, E. J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic Bulletin & Review, 14*, 779–804.

Wilson, T. D., Laser, P. S., & Stone, J. I. (1982). Judging the predictors of one's own mood: Accuracy and the use of shared theories. *Journal of Experimental Social Psychology, 18*, 537–556.

Wood, D., & Brumbaugh, C. C. (2009). Using revealed mate preferences to evaluate market force and differential preference explanations for mate selection. *Journal of Personality and Social Psychology, 96*, 1226–1244.

Wood, D., & Furr, R. M. (2016). The correlates of similarity estimates are often misleadingly positive: The nature and scope of the problem, and some solutions. *Personality and Social Psychology Review, 20*, 79–99.

Wood, D., Lowman, G. H., Harms, P. D., & Roberts, B. W. (2019). Exploring the relative importance of normative and distinctive organizational preferences as predictors of work attitudes. *Journal of Applied Psychology, 104*, 270–292.