Consistency and Inconsistency among Romantic Partners over Time

Paul W. Eastwick[1]

K. Paige Harden[2,3]

Jennifer A. Shukusky[4]

Taylor Anne Morgan[5]

Samantha Joel[6]

In Press, *Journal of Personality and Social Psychology*

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

[2] Department of Psychology, University of Texas at Austin

[3] Population Research Center, University of Texas at Austin

[4] Department of Human Development and Family Sciences, University of Texas at Austin

[5] Facebook, Menlo Park, California

[6] Department of Psychology, University of Utah

Paul W. Eastwick
University of California, Davis
Department of Psychology
One Shields Ave
Davis, CA 95616
Tel: (773) 484 3878
Email: eastwick@ucdavis.edu

Abstract

Theoretical perspectives on mating differentially emphasize whether (and why) romantic partner selection and maintenance processes derive from stable features of individuals (e.g., mate value, mate preferences, relationship aptitude) and their environments (e.g., social homogamy) rather than adventitious, dyad-specific, or unpredictable factors. The current article advances our understanding of this issue by assessing how people's actual romantic partners vary on constructs commonly assessed in evolutionary psychology (Study 1), sociology (Study 2), and relationship science (Study 3). Specifically, we calculated the extent to which the past and present partners of a focal person (i.e., the person who dated all of the partners) *cluster* on various measures. Study 1 investigated consistency in the observable qualities of the romantic partners, revealing substantial evidence for clustering on coder-rated attributes like attractiveness and masculinity. Study 2 examined qualities self-reported by romantic partners themselves in a demographically diverse sample and found modest evidence for clustering on attributes such as IQ and educational aspirations; however, clustering in this study was largely due to demographic stratification. Study 3 explored target-specific ratings by partners about the focal person and found little evidence for clustering: The ability to elicit high romantic desirability/sexual satisfaction ratings from partners was not a stable individual difference. The variables that affect mating may differ considerably in the extent to which they serve as stable vs. unpredictable factors; thus, the fields of evolutionary psychology, sociology, and relationship science may reveal distinct depictions of mating because the constructs and assessment strategies in each differ along this underappreciated dimension.

Keywords: Human mating, mate selection, attraction, romantic relationships

Consistency and Inconsistency among Romantic Partners over Time

Researchers across several scientific literatures examine how humans select one another as mates and maintain those mateships over time. However, the disciplines' depictions of this process do not always cleanly intersect; it would not be unusual to encounter a few papers from each of the evolutionary psychological, sociological, and close relationships literatures and marvel at how they all describe mating in a single species rather than two or three different species. Is *Homo sapiens* a species that exhibits pronounced observable differences in romantic desirability, as emphasized by evolutionary psychological perspectives on mate value (Kavanagh, Robins, & Ellis, 2010; Kenrick, Trost, Groth, & Sadalla, 1993; Penke, Todd, Lenton, & Fasolo, 2007), a species whose partner choices are shaped by economic and demographic contexts, as emphasized by sociological perspectives on assortative mating (Burgess & Wallin, 1953; Kalmijn, 1998; Schwartz & Mare, 2012), or a species in which mating decisions are driven by subjectively biased perceptions of a specific partner, as emphasized by close relationships perspectives on relationship maintenance (Le, Dove, Agnew, Korn, & Mutso, 2010; Murray, Holmes, & Griffin, 1996; Rusbult, Olson, Davis, & Hannon, 2001)? The clear answer is "all of the above" and will require an extensive integrative effort.

In most cases, the scientific conclusions that follow from evolutionary, sociological, and close relationships perspectives may all be correct yet nonintersecting because the disciplines (a) study different constructs and (b) use different analytical approaches. For example, common methodological approaches include the study of observable, desirable qualities (e.g., attractiveness) using photographs in the evolutionary psychological literature (Little, Jones, & DeBruine, 2011), the study of demographically variable constructs (e.g., education level) using self-report surveys or interviews in the sociological literature (Schwartz & Mare, 2012), and the

study of people's subjective reports (e.g., relationship satisfaction) about a specific other person, usually a romantic partner, in the close relationships literature (Fletcher, Simpson, & Thomas, 2000b). Given these methodological choices, it should come as little surprise that the depictions of mate selection processes in the three disciplines emphasize observable desirability, demographic sorting, and partner-specific biases, respectively. Furthermore, these methods may be driving theory rather than vice versa, because depending on which literature one peruses, it would be easy to conclude from prominent theories that human mating is fundamentally located in stable features of the individual, stable features of the environment, or time- and context-sensitive features of dyads.

One approach that may aid in the long-term goal of integrating these perspectives is to use a single analytic strategy that can be applied to any and all of these different types of variables—that is, one could study the constructs common in all three literatures but hold constant the data analytic approach. This manuscript pioneers the use of such a data analytic strategy: calculating the extent to which a person's current and former romantic partners *cluster* (i.e., are more similar to one another than would be expected due to chance) on particular variables of interest. We use this strategy to address a substantive question inherent to all three disciplines: To what extent are a person's romantic choices driven by stable features of the chooser and his/her environment rather than factors that shift from relationship to relationship or are fundamentally unpredictable (Lykken & Tellegen, 1993)? We raise this question not in an attempt to garner support for the theoretical perspectives of one literature over another. Rather we assume that scientists want to explain how people choose romantic partners in the real world, and therefore they endeavor to understand how observable features, demographics, and relationship-specific judgments factor into this process. Given this goal, the extent of clustering

offers useful clues about where the largest causal forces are likely to reside (e.g., stable vs. unstable factors) for a given construct of interest.

Datasets addressing the extent of clustering in partner choices over time are rare and challenging to collect. Whereas dyadic data are common in the close relationships field, dyadic data deriving from a given focal person's *multiple* romantic partnerships are nearly nonexistent (for one exception, see Robins, Caspi, & Moffitt, 2002). The current manuscript contains three studies that overcome this hurdle, and each examines constructs that largely correspond to one of the three literatures. Study 1 finds moderate-to-strong evidence for clustering on immediately observable variables (e.g., attractiveness) common in the evolutionary psychological literature. Study 2 finds qualified evidence for clustering on demographic variables (e.g. educational aspirations) common in the sociological literature. Finally, Study 3 finds little evidence for clustering on—and thus little evidence that stable forces affect—partner-specific judgments (e.g., romantic desire reports about a partner) that are common in the close relationships literature. We also present a simulation study demonstrating that, in principle, clustering can emerge from active partner selection processes alone, even when the number of observed partners in the dataset is small (i.e., two partners per focal person). We conclude with a discussion of how the current results and similar investigations might help to unify these different perspectives on mating.

Stable Influences on Romantic Partner Clustering

The people with whom we *could* form relationships might differ from the people with whom we *do* form relationships, and this comparison can reveal important insights about the processes underlying human mating. If people selected romantic partners at random, then the current and former partners of a given focal person would be no more similar to one another than

to any other individual in the population. In contrast, if predictable, stable factors bring dyads together beyond chance alone, a focal person's romantic partners will cluster, which means that they will exhibit similarities that are not shared with other individuals who have never been the focal person's relationship partner. For example, nonrandom, stable factors such as a focal person's own attributes (e.g., their own intelligence), mate preferences (e.g., a preference for intelligent partners), and consistent environments (e.g., living near intelligent people) will cause a focal person's current and former romantic partners to possess similar attributes (e.g., high intelligence). These same forces could also cause clustering in the ratings provided by the current and former partners about the focal person: A focal person's desirable qualities should cause his/her partners to provide similar ratings about his/her romantic desirability. The extent to which current and former romantic partners cluster on a given attribute, therefore, denotes the upper limit on the extent to which factors that are consistent across time (e.g., a focal person's stable attributes, personality, or living context) influence mating with respect to that attribute.[1]

Active stable factors. Many prominent theories of human mating predict that people's stable qualities cause them to select some partners over others—a process that should produce clustering. Furthermore, much of this selection process is presumed to be active, which means that it emerges as individuals (a) evaluate and/or (b) are evaluated by potential partners in their immediate social milieu. Several examples of these theories follow: Firstly, evolutionary perspectives suggest that some people have greater *mate value* than others (i.e., they possess traits that would promote a partner's reproductive success; Sugiyama, 2005), and people high in mate value should be better than people low in mate value at attracting and retaining partners with desirable qualities (Ellis & Kelley, 1999; Kirsner, Figueredo, & Jacobs, 2003; Penke et al., 2007). In other words, individual differences in a focal person's mate value should cause

clustering among current and former partners in terms of the presence or absence of normatively desirable traits, such as physical attractiveness, appealing personality traits, or intelligence. Secondly, evolutionary and close relationships perspectives suggest that people also possess idiosyncratic *mate preferences* for particular qualities in romantic partners—qualities that are desirable to some people but not to others (Buss & Barnes, 1986; Fletcher, Simpson, Thomas, & Giles, 1999). If people select partners who are congruent with their mate preferences, then clustering should also emerge for idiosyncratically desirable traits: For example, people with a preference for energetic, adventurous partners should be more likely than people without this preference to date energetic, adventurous partners. Thirdly, *similarity-attraction* effects (Buston & Emlen, 2003; Byrne, 1961; Montoya, Horton, and Kirchner, 2008) should also produce clustering among current and former partners: Even if people are unaware whether they do or do not have preference for religiosity, clustering on religiosity will emerge if people tend to be drawn to similar others. Finally, in the relationship maintenance domain, some people might have greater *relationship aptitude* than others as a function of their particular personality traits or personal histories (e.g., neuroticism, attachment anxiety), and this aptitude would cause clustering to emerge for partners' target-specific judgments of a focal person (Finkel, Eastwick, Karney, Reis, & Sprecher, 2012). For example, some people may be better able to resolve conflict and respond empathically to their partners, and these abilities should in turn make their partners similarly satisfied or unsatisfied in their relationship with the person.

No prior studies have assessed clustering among a single focal person's multiple romantic partners over time; nevertheless, the existence of assortative mating strongly implies that clustering should occur. For example, the two members of a romantic pair tend to sort with respect to appearance—most notably physical attractiveness (Feingold, 1988)—as well as other

desirable qualities like intelligence (Watson et al., 2004). Researchers continue to investigate whether factors like mate value or mate preferences can explain the extent to which partners sort on these qualities (Burriss, Welling, & Puts, 2011a, 2011b; Conroy-Beam & Buss, 2016; DeBruine et al., 2006; Hunt, Eastwick, & Finkel, 2015; Kalick & Hamilton, 1986). Yet as long as the underlying causal factors, like mate value, exhibit at least modest stability as people move from relationship to relationship, then the same factors that produce assortative mating should produce clustering in a person's romantic partners over time. In other words, if (a) a man's observable mate value when he selects partner A correlates positively with his observable mate value when he later selects partner B, and (b) mate value produces assortative mating (i.e., high mate value individuals successfully attract other high mate value individuals), then (c) partner A and partner B should cluster on mate value.

Passive stable factors. Even in the absence of an active psychological selection process, passive yet stable factors—factors that affect which dating partners a focal person has the opportunity to meet—may also produce clustering. For example, *social homogamy* perspectives note that some portion of the mate selection process originates in the social milieu that surrounds each individual. People are more likely to meet and date others who live nearby and whom they encounter frequently (Belot & Francesconi, 2013; Festinger, Schacter, & Back, 1950; Luo & Klohnen, 2005; Newcomb, 1961), and people tend to live near others who are similar to them with regard to income, educational attainment, and race. The schools that adolescents and young adults attend are especially important sorting factors in this regard (Mare, 1991). Because young people's acquaintances are highly likely to attend the same school, the extent to which attributes cluster at the school level should also affect the clustering of acquaintances and, subsequently, clustering of romantic partners. In other words, the tendency for people to meet others who are

similar to them by virtue of their living situation will produce clustering among romantic partners (Mascie-Taylor & Vandenberg, 1988). Although social homogamy forces are relatively passive and do not imply that a focal person exerts any predictable choice of romantic partners *within* his or her immediately available pool of partners, they may nevertheless have large effects on producing clustering. Indeed, just as with the active forces described above, the existence of assortative mating on variables that differ across demographic contexts (e.g., education, religiosity; Watson et al., 2004) strongly implies that clustering should emerge among a focal persons' current and past romantic partners with respect to these variables.

Adventitious Perspectives on Romantic Partner Clustering

For many constructs, clustering among romantic partners could be quite substantial. But other perspectives emphasize how much of the mate selection process in humans may be driven by chance, unpredictable forces not rooted in stable characteristics of individuals or their environments. In some contexts, the active selection factors described above may be quite weak. For example, once a face-to-face interaction has occurred, there is no replicable evidence that people are more likely to select mates who match rather than mismatch their preferences for a particular attribute (Eastwick et al., 2014a; cf. Campbell, Chin, & Stanton, 2016). With respect to mate value, agreement among opposite-sex acquaintances in terms of who does and does not possess mate value is extremely low, so consensual mate value may not have a strong influence on mate selection in the case where romantic partners get to know each another before forming a relationship (Eastwick & Hunt, 2014). Finally, the similarity-attraction effect tends to be weak in face-to-face initial attraction settings (Luo & Zhang, 2009), as well as in established relationships (Watson et al., 2004). Although stable forces such as mate preferences and mate value are surely influential under some circumstances (e.g., when perusing online dating profiles;

Eastwick et al., 2014a), the extent to which they account for variance in determining whom people actually select as romantic partners remains unclear.

These alternative perspectives on partner selection do not imply that the underlying process is atheoretical and random. Rather, stable influences on romantic outcomes will appear to be weak to the extent that romantic selection processes are highly dyadic, synergistic, or contextual. For example, mate preferences for traits may not generate clustering because traits change their meaning depending on the context of a partner's overall constellation of traits (Eastwick, Finkel, & Eagly, 2011). Akin to context effects in classic person perception research (Asch, 1946), people may not consistently select romantic partners who have a particular trait because a trait that that makes one partner appear desirable may make another partner seem undesirable. Furthermore, mate value judgments consist of large amounts of idiosyncratic variance (i.e., relationship variance; Eastwick & Hunt, 2014), and thus a wonderful partner for one person might be a terrible partner for another person, irrespective of consensual mate value. Myriad theories of close relationships are consistent with this conceptualization of the mate selection process—theories that highlight the importance of the dyad (e.g., Kelley et al., 2003), the way that relationships grow and change over time (e.g., Hazan & Shaver, 1994; Karney & Bradbury, 1995), and how idiosyncratic factors have large influences on relationship outcomes and yet are largely unknowable before two people meet and get to know each other (Eastwick, 2016; Eastwick, Keneski, Morgan, & MacDonald, 2016; Finkel et al., 2012). Many complex social processes can be predicted only weakly from distal predictors alone (e.g., economic fundamentals poorly predict election outcomes; Lauderdale & Linzer, 2015; Silver, 2012); romantic partner selection and maintenance processes may be similarly difficult to predict from trait information and other self-reported qualities.

Another source of data suggesting that human mating is largely governed by adventitious factors derives from twin data (Lykken & Tellegen, 1993). Although many qualities of the spouses of twins are correlated (e.g., the church activities of co-twins' spouses correlate at approximately $r = .30$), these correlations do not differ between monozygotic twins and dizygotic twins. This pattern stands in marked contrast to other preferences and choices (e.g., jobs, leisure activities, friends), which typically show greater correlations among monozygotic than dizygotic twins (Bouchard, Lykken, McGue, Segal, & Tellegen, 1990). In other words, the lack of a difference between monozygotic and dizygotic twins' partner choices (as evidenced by their similarity levels) suggests that the internal, genetically influenced processes that affect many different kinds of life choices apply only weakly to the choice of a mate (see also Zietsch, Verweij, Heath, & Martin, 2011). Informed by these findings, Lykken and Tellegen (1993) proposed that successful reproduction in ancestral environments primarily depended on forming and maintaining a strong pair-bond, not selecting the best possible mate (see also Hazan & Diamond, 2000).

The Current Research

The similarity inherent to a person's unique pool of romantic partners, past and present, denotes the extent to which mate selection on a given attribute is governed by stable, predictable factors, both active (e.g., mate preferences) and passive (e.g., social stratification). Surprisingly, the extent of similarity has not been previously calculated with respect to *any* attribute, perhaps because of the challenges inherent in an analytic strategy that requires information provided by multiple romantic partners. One study examined participants' *own* reports of relationship experiences with two different partners at two different points in time (Robins et al., 2002), but

partners were themselves recruited to participate at only one time point, so no self-reported or objective qualities of the multiple partners were available to calculate clustering.

The current set of studies addresses this critical gap in the mate selection literature, and it represents an initial attempt to quantify the extent to which partner choice is predictable with respect to several different romantic partner attributes. First, a Simulation Study uses agent-based modeling in order to test whether active mate selection strategies would result in empirically-detectable clustering in data sets that sampled a restricted number of mates (2-4) for each individual. Next, we present results from three empirical studies. Study 1 examines clustering among focal persons' current and former partners in terms of traits that are observable from the partners' photographs. After all, people can tell a great deal about another person from a photograph, and evolutionary psychologists have capitalized on this feature in their study designs, which have frequently emphasized desirable traits such as attractiveness, masculinity, and dominance (e.g., Bailey, Durante, & Geary, 2011; Burriss et al., 2011a, 2011b; Fink & Penton-Voak, 2002; Little et al., 2011; Maner, Dewall, & Gailliot, 2008; Rhodes, 2006; Zietsch, Lee, Sherlock, & Jern, 2015). Study 2 uses data from a large, nationally representative, longitudinal study of adolescents and young adults: the National Longitudinal Survey of Adolescent Health (Add Health; Harris & Udry, 1994-2008). This survey included several traits related to romantic desirability that were self-reported by the current and former partners of focal persons, and because some of the traits vary across demographic contexts (e.g., IQ), we could quantify the extent to which clustering for these variables is a function of active vs. passive selection forces. Study 3 addresses whether clustering emerges with respect to target-specific reports—the most common assessment strategy in the close relationships literature (Fletcher et

al., 2000b)—by drawing from an internet database containing women's ratings of men whom they were currently dating or had dated in the past.

In all three studies, the datasets can be organized such that romantic partners are nested within focal person—the participant who chose to date all of those partners. The intraclass correlations (ICC) for the focal person capture the extent to which qualities of romantic partners cluster. For example, the focal person intraclass correlation for intelligence indicates the extent to which focal persons consistently dated intelligent versus unintelligent partners across time. This consistency across current and former partners could be due to the focal person's mate value (possessing normatively desirable characteristics that make one better able to attract intelligent partners), the focal person's mate preferences (prioritizing intelligence over other attributes in mating decisions), similarity-attraction (intelligent people are more likely to date other intelligent people), or social homogamy (a person's pool of potential mates is drawn from an educational context that is stratified by intelligence). With respect to target-specific reports (Study 3), consistency across current and former partners could also be due to the focal person's relationship aptitude (some focal persons have relationship skills that cause their partners to rate them as desirable; Finkel et al., 2012).

As the discussion above highlights, clustering should emerge on variables that people generally want in their partners (e.g., any construct that taps mate value or relationship aptitude), variables that some people want more than others (e.g., any construct that exhibits idiosyncratic differences in mate preferences), or variables on which people sort based on location (e.g., any construct linked to social homogamy). All of the constructs examined in the present set of studies are linked to at least one of these perspectives—that is, each variable (a) is desired by people in general, (b) is desired by some people more than others, and/or (c) is related to demographic

sorting. Furthermore, each variable has strong theoretical connections to one or more of the evolutionary psychological, sociological, and close relationships perspectives, and each has been examined frequently in a mating-relevant context using the same assessment strategy (e.g., rated from a photograph, self-reported) in prior research (except the number of positive/negative qualities variables in Study 3, see Supplementary Table 1).

Simulation Study

In this simulation study, we sought to ensure that our data analytic strategy is, in fact, able to detect effects of stable influences on romantic clustering. In the studies that follow, the average number of partners per focal person is approximately four in Study 1, two in Study 2, and three in Study 3. It is unclear whether clustering for attributes could emerge in a dataset that includes a sample of targets in this range, even if people select partners actively on the basis of their mate preferences for those attributes. To examine the extent to which clustering is empirically detectable, we used *agent-based modeling*, which involves simulating environments in which hypothetical human actors behave according to theoretically relevant behavioral rules (Conroy-Beam & Buss, 2016; Smith & Conrey, 2007). Specifically, we conducted a series of computer simulations in which "agents" selected "mates" based on the match between the agents' preferences and the mates' traits.[2] These simulations test the following question: If we consider even just a single stable influence (i.e., selection based on mate preferences), setting aside the many other possible stable influences that could exist in the real world (e.g., mate value, similarity-attraction, social homogamy), could a detectable amount of clustering emerge in a set of 2-4 selected mates?

Method

We constructed an agent-based model designed to represent an environment in which focal persons select mates based entirely on the match between (a) the focal person's preferences for particular attributes and (b) the extent to which they perceive that their available mates possess those attributes. For the purposes of this simulation, an agent represents a focal person seeking a romantic partner, and the mates are the partners available for selection by the agent. Analyses were conducted in R (see R code here).

Model Assumptions

One critical assumption that we incorporated into the model concerns the underlying structure of the agent's preferences and the agent's perceptions of the partners' traits. In real life, these judgments correlate: People who rate attractiveness highly in an ideal partner are also more likely to rate intelligence highly, and people who perceive their partners as being attractive also perceive them to be intelligent. We conducted two sets of simulations to reflect two different ways that humans' preference and perceptual architecture might be structured. In one set of simulations ("modest intercorrelations"), ideal partner preferences ratings correlated with other ideal partner preference ratings at $r = .20$ and partner trait ratings correlated with other partner traits ratings at $r = .20$. In a second set of simulations ("stronger intercorrelations"), ideal partner preferences ratings correlated with other ideal partner preference ratings at $r = .20$ and partner trait ratings correlated with other partner traits ratings at $r = .45$ (see Supplementary Materials for justification of these values).

A second assumption concerns the number of preferences that the agents possessed (and the corresponding number of traits on which the mates varied). It is unclear how many trait dimensions people use to evaluate romantic partners in real life. The most intensively validated ideal partner preference measure contains three (correlated) factors (warmth/trustworthiness,

attractiveness/vitality, and status/resources; Fletcher et al., 1999). Marlowe (2004) used a method similar to the one used by Fletcher et al. (1999) with a sample of Hadza hunter-gatherers and documented seven factors, and other sets of preference ratings have revealed seven factors as well (e.g., Eastwick et al., 2011, Study 3). Complicating the picture further is that people may vary in the extent to which they *use* different trait dimensions to evaluate partners; that is, even if the true number of trait factors that exist in the world is three or seven or even higher, people may not compare partners to their ideals on all dimensions. To address this complex issue, we conducted several simulations using agents that possessed between 1 and 20 preferences (in increments of one).

Agent-based Simulation

Each cycle of the model generated 100 agents with 1 to 20 preferences that were pre-specified to correlate with each other at $r = .20$. Then, a total of 100 potential mates were generated for each agent with the same number of traits (i.e., between 1 and 20), and the traits were also prespecified to correlate with each other at $r = .20$ ("modest intercorrelations") or $r = .45$ ("stronger intercorrelations"). (No correlation was prespecified between the preferences of the agents and the traits of the mates.) The values for each preference and trait were drawn from a multivariate normal distribution, and all traits and preferences had a pre-specified mean of 4. The degree to which each agent was "attracted" to each of their 100 potential mates was determined by computing the sum of the squared deviations between the agent's preferences and the potential mate's traits (Conroy-Beam & Buss, 2016).

Each agent then selected their two, three, or four most attractive mates (in three separate sets of simulations). We chose these values to correspond with the average number of mates that we happened to acquire in the datasets corresponding to Studies 2, 3, and 1, respectively, in this

article. We then calculated the ICC for each trait across the 2-4 mates selected by each agent in each cycle. These ICC values represent how similar a person's romantic partners would be (i.e., extent of clustering) with respect to a particular attribute, in principle, in an environment in which each person consistently chooses 2-4 mates according to their own idiosyncratic mate preferences. Typically, this variance estimate is considered to be "meaningful" if it reaches at least r = .10 (Kenny, Kashy, & Cook, 2006); .20 is a reasonable benchmark for a medium-sized effect and .30 is a reasonable benchmark for a large effect (see also Gignac & Szodorai, 2016). For all simulations that contained between 2 and 20 traits, the ICCs were averaged across the 2-20 traits in the figures below.

Results

Figure 1 depicts the mean ICCs across the 2-4 mates as a function of the number of preferences/traits included in each cycle. Panels A-C present findings for agents selecting their top 2, 3, and 4 mates, respectively, although findings for the three panels are largely similar. Not surprisingly, when there was only one trait on which agents were evaluating mates, the ICC was extremely close to 1.0: That is, clustering on the selected trait was nearly perfect, which is not surprising because the agents were designed to select mates based on the (one) trait and because there were no constraints on the agents' ability to mate with preferred partners. As the number of traits increased, the ICC values fell but still remained fairly large: ICCs were above .80 in all three graphs if agents evaluated mates based on three trait constructs (as suggested by Fletcher et al., 1999), and it was in the .50-.60 range if agents evaluated mates based on seven constructs (as suggested by Marlowe, 2004, and Eastwick et al., 2011, Study 3). As the simulations contain more and more traits, the ICC fell further, but it began to reach an asymptote around ICC = .30. In other words, in a world where people select mates solely based on the extent to which mates

match their preferences, the ICC for a given trait in a pool of people's selected mates is at least medium-sized on average, even if only a small number of selected mates are contained in the dataset.

The "modest intercorrelations" estimates assumed that preferences correlate with other preferences at $r = .20$ and that perceptions of a mate's traits correlate with other traits at $r = .20$. For the "stronger intercorrelations" simulations, we replaced the $r = .20$ correlation among traits with values that may more appropriately reflect an initial attraction context (i.e., $r = .45$; Eastwick & Finkel, 2008), and the ICCs increased somewhat. The difference was small when the number of traits was small but grows as the number of traits increases, approaching an approximate asymptote of ICC = .40. In summary, in an initial selection context that contained a more pronounced correlation between the perceived traits of the potential mates, ICCs could be quite large in principle. (See Supplementary Materials for additional simulations incorporating reliability of measurement.)

Discussion

These simulations suggest that, in a world where people select mates based on their preferences for particular attributes, people's romantic partners would cluster on those attributes. ICCs were substantial even when only two partners were actually selected, and they remained medium-sized even when people selected mates based on 20 different attributes—a number that is quite large relative to previous well-validated research on the number of constructs that people consider in mate selection contexts (e.g., Fletcher et al., 1999).

One important element that we did not incorporate into these simulations was the reciprocal nature of partner choice (Conroy-Beam & Buss, 2016); in real life, partners select each other. A more externally valid context that involved reciprocal choice could affect ICCs

both positively and negatively in principle. If the agents were willing to partner up with (for example) the top 10 mates to whom they were attracted, and reciprocal liking was random within that subset of sufficiently desirable mates, then ICCs would surely decrease. On the other hand, as other stable forces like mate value come into play, highly desirable individuals should be more likely to select each other as partners, which would increase ICCs. In other words, the reciprocal nature of mating could function as either a stable factor, thereby increasing the ICC, or an unstable factor, thereby decreasing the ICC. Future simulations could expand the realism of the initial simulations we report here. For the present purposes, the most important element of this simulation study is that, even in the presence of a single force operating as a stable influence on mate selection (e.g., selecting mates based on the extent to which they match preferences), substantial ICCs can emerge even when the number of selected mates is small (i.e., two).

Study 1

People possess mental mechanisms that evolved to guide attention to and selectively process information about potential mates based on the visible attributes that those potential mates possess (Maner et al., 2008; Maner, Gailliot, & DeWall, 2007). Consistent with this perspective, the large person-perception literature on face perception has revealed how people routinely extract a great deal of information from faces (Oosterhof & Todorov, 2008; Rule, Ambady, Adams, & McCrae, 2008). Indeed, many studies in the evolutionary psychological tradition have drawn from these theories to explore how people behave in response to faces that vary in attractiveness, masculinity, and dominance (for examples, see Supplementary Table 1). These observable factors are likely to matter early in the mate selection process as people attempt to identify potential mating opportunities (Maner et al., 2007; 2008).

Building on this literature, Study 1 examined the extent to which a person's actual partners cluster on these observable attributes. If stable forces such as mate preferences, mate

value, or similarity-attraction operate with respect to these sorts of observable attributes when people are initially selecting mates (e.g., people with a preference for these qualities select partners who possess them; people who have the qualities can attract partners who possess them), then people's past and current partners should exhibit clustering on these features (e.g., some people should have more attractive partners than others). Importantly, the extent of clustering on each attribute is an indicator of the strength of all stable forces over time with respect to that attribute.

Method

Participants

These data were taken from a study of 136 university students who completed a 30-minute study in the laboratory for course credit. Our goal was to collect at least 100 focal persons for this study during a single academic semester. We collected more participants than this target because we anticipated that (a) some participants would not be able to (or want to) report on at least two current/ex romantic partners (N = 22), and (b) some participants would not be able to provide usable photographs for at least two current/ex romantic partners (N = 15). Also, although we initially planned to analyze all eligible participants' data regardless of sexual orientation, the inclusion of the N = 2 participants who nominated same-sex partners affected the findings dramatically, especially on the masculinity variable (i.e., the inclusion of the two women who nominated female partners boosted the appearance of clustering on masculinity substantially, probably because their partners—women—were markedly less masculine than other female participants' partners—men). Thus, we do not include these N = 2 participants in the analyses.

In summary, the 136 participants in the study gave us a usable sample of N = 97 (12 men, 85 women) participants who provided internet links to usable photographs for two or more

current/ex romantic partners; analyses below were conducted on this subsample. These 97 focal persons were 20.0 years old on average (*SD* = 1.2, range = 18-24); 8.2% identified as Black, African American, Caribbean American, 22.7% as Asian-American, Asian, Pacific Islander; 38.1% as White, European-American, Anglo, Caucasian; 24.7% as Hispanic-American, Latino(a), Chicano(a); 4.1% as Bi-racial, Multi-racial; and 2.1% as Other. The average response to the item "I am exclusively attracted to members of the opposite-sex" was *M* = 8.4 (*SD* = 1.4) on a scale from 1 (*very strongly disagree*) to 9 (*very strongly agree*).

Procedure and Materials

Participants (i.e., focal persons) arrived at the laboratory and were greeted by an experimenter. The participant learned that the study examined how people depict themselves on social media, and he/she would be asked to provide links to publicly available photographs (i.e., the Facebook profile photograph) of several friends and acquaintances. The experimenter then demonstrated how to copy and paste the url for a Facebook profile photograph (which does not include identifying information) into the survey text boxes.

At the beginning of the survey, participants first provided "the first name and last initial of the last 8 people with whom you have engaged in sexual behavior (THIS MAY ONLY INCLUDE KISSING, but may also include oral sex, sexual intercourse, etc.)." They were instructed to include each person only once and to enter personally meaningful descriptors (e.g., guy from party, friend's roommate) if they could not remember the partner's name. Also, participants were instructed to leave the corresponding spaces blank if they possessed fewer than eight previous partners. The 97 focal persons listed a total of 553 *romantic partners*. (Participants then provided the names of platonic friends and crushes; these data are not analyzed in the current study.)

On subsequent pages of the survey, participants provided urls for the Facebook profile photograph for each of the romantic partners they had listed earlier. Of the 553 listed partners, participants provided usable links to 428 of them (M = 4.4 partners per participant, SD = 2.0; range = 2-8; as noted above, all 97 participants included in the analyses provided at least two usable photographs, otherwise the intraclass correlation cannot be calculated). If the photograph displayed multiple people, the participants indicated which person was the partner in a text box. When photographs were unavailable (N = 125), participants gave one of the following reasons: the partner had blocked them on Facebook (N = 5), he/she had blocked the partner (N = 10), the partner did not have Facebook (N = 48), and "Other" or no reason provided (N = 51). Finally, participants provided links for N = 11 partners where the link was either broken or the photograph did not clearly depict the partner. Participants also indicated for each partner whether he/she was (a) a "current boyfriend/girlfriend" (N = 45), (b) an "ex/former boyfriend/girlfriend" (N = 128), (c) a "current non-committed partner (e.g., hookup, friend with benefits)" (N = 24), or (d) a "past/former non-committed partner (e.g., hookup, friend with benefits)" (N = 231). At the end of the survey, 95 of the 97 participants also provided links to usable Facebook photographs of themselves (i.e., focal person photographs).

At the end of the academic semester (i.e., after all participants had completed the survey), we downloaded the photographs, edited them to be 300 pixels high and 200-300 pixels wide, and blurred or cropped out other people besides the partner. Then, ten research assistants (seven women and three men) rated all usable photographs (presented randomly); the research assistants came from the same student population as the focal persons and were similarly racially/ethnically diverse (20% identified as Black, African American, Caribbean American, 30% as Asian-American, Asian, Pacific Islander; 30% as White, European-American, Anglo,

Caucasian; and 20% as Bi-racial, Multi-racial). Using scales from 1 (*not at all*) to 9 (*a great deal*), the research assistants rated the photographs on the following three constructs: *Attractiveness* ("physically attractive", "sexy/hot"; α = .96 across the two items for partner photographs and α = .98 across the two items for focal person photographs), *masculinity* ("masculine", "feminine" [reverse-scored]; partner photograph α = .97 and focal person photograph α = .97), and *dominance* ("dominant", "confident"; partner photograph α = .95 and focal person photograph α = .86). Agreement across the ten research assistants was strong for ratings of partner photographs (attractiveness α = .82, masculinity α = .96, dominance α = .81) and for focal person photographs (attractiveness α = .90, masculinity α = .96, dominance α = .77).

Analysis Strategy

Using PROC MIXED in SAS, we calculated the focal person intraclass correlation (ICC) for each of the three constructs across all partner photograph ratings; this metric is conceptually identical to the percentage of variance accounted for by the focal person. In the dataset, each row consisted of the ratings of a single partner photograph, and an indicator variable (i.e., focal person ID) linked each partner to his/her focal person. We treated focal person ID as a random factor, and the statistic of interest is the random variance estimate for the focal person ID divided by the total variance (i.e., the percentage of variance in partners' characteristics due to the focal person). This estimate indicates the extent to which partners with a particular characteristic (i.e., attractiveness, masculinity, dominance) are more likely to be clustered around some focal persons rather than others. Once again, this variance estimate is "meaningful" if it reaches at least 10% (i.e., an ICC of r = .10; Kenny et al., 2006); 20% is a reasonable benchmark for a medium-sized effect and 30% is a reasonable benchmark for a large effect.

Men and women are likely to differ on several of the characteristics examined in this study, and these sex differences could lead to the illusion of clustering. For example, because some focal persons have male partners and others have female partners, for this reason alone the partners of some focal persons (i.e., women) will presumably be more masculine than the partners of other focal persons (i.e., men). Thus, all analyses control for the sex of the focal person.

Several theories of mating suggest that people desire different types of partners for long-term vs. short-term relationships (Buss & Schmitt, 1993; Gangestad & Simpson, 2000), and it is possible that people are more selective when choosing long-term than short-term partners. Thus, clustering might be weakened by the inclusion of the non-committed partners along with the boyfriend/girlfriend partners in the analyses. In order to address this possibility, subsidiary analyses below include only the boyfriend/girlfriend partners—not the non-committed partners—to see if the results change appreciably. Data and code for Study 1 are available here.

Results

Main Analyses

Table 1 presents the percentage of variance accounted for by focal person across the three photo-rated qualities. The percentage of variance due to the focal person was generally moderate to large, ranging from 27% (attractiveness) to 31% (masculinity); the average variance across all three qualities was 29.1%, and all were significantly greater than zero. In other words, there was substantial clustering of partner qualities by focal person: Some people tended to have romantic partners who were attractive or masculine or dominant, whereas other people tended to have romantic partners with low levels of these qualities. Recall that these analyses control for participant sex, so the fact that female partners were rated by the research assistants as more

attractive than male partners (M_{female} = 4.76, SD_{female} = 1.16, M_{male} = 3.45, SD_{male} = 0.88, d = 1.43) or that male partners were rated as more masculine than female partners (M_{female} = 2.43, SD_{female} = 0.56, M_{male} = 6.85, SD_{male} = 0.74, d = 6.12) cannot account for these findings.

We also examined whether the percentages in Table 1 significantly differed for men and women: That is, were the partners of male focal persons more likely to cluster on a given variable than the partners of female focal persons? Of the three variances, one marginally significantly differed by sex: Female focal persons were more likely to exhibit clustering for masculinity (31.8%) than male focal persons (15.2%), z = 1.89, p = .059. In other words, the male partners of a given female focal person exhibited strong similarity in the extent to which they were masculine versus feminine, whereas the female partners of given male focal person exhibited only modest similarity in the extent to which they were masculine versus feminine.

The research assistants also rated the focal persons themselves on the three qualities, so we could therefore examine the extent to which similarity played a role in predicting clustering. Similarity effects are presented in the second column of Table 1; these values indicate the standardized beta for the focal person's score on the attribute predicting the partners' scores on the attribute (as a fixed effect). All three effects were small-to-moderate in size, indicating that attractive focal persons tended to have attractive partners, masculine focal persons tended to have feminine partners, and dominant focal persons tended to have dominant partners. Controlling for similarity caused variance due to the focal person to decline, as expected; the extent of reduction ranged from two to eight percent.

Subsidiary Analyses

In principle, assortative mating on race or ethnicity could cause clustering on attractiveness, masculinity, or dominance to emerge. If (a) focal persons tended to date partners

who were of a similar race/ethnicity, and (b) racial/ethnic groups received different ratings on any of the three qualities, then (c) clustering could emerge due to racial/ethnic sorting alone. Unfortunately, we did not possess race/ethnicity information about the targets; thus, we cannot directly address this question. Nevertheless, focal persons did self-report their race/ethnicity, and so we can include focal person race as a categorical fixed effect in the analyses. Including focal person race caused the level of clustering to change very little: clustering in these three analyses was 27.9% for attractiveness, 21.3% for masculinity, and 26.3% for dominance, and all three remained significantly different from zero. Furthermore, we conducted the three analyses separately on the three racial/ethnic groups for which we possessed data from at least $N = 20$ focal persons (i.e., Asian-American, White, and Hispanic-American focal persons), and the degree of clustering averaged 30.8% and was significantly or marginally significantly greater than zero in eight out of nine tests. These analyses suggest that it is unlikely that clustering in this study was due to assortative mating by race or ethnicity. (We directly examine clustering by target race in Study 2 because targets self-reported their race in that study.)

Also, assortative mating based on the age of the focal person could serve as an alternative explanation for the clustering we observed in this study. For example, given the well-established negative association between age and coder-ratings of physical attractiveness (e.g., Perrett et al., 2002), we might see clustering emerge for these qualities simply because focal persons are likely to date similarly aged partners. However, controlling for age as a fixed effect changed the findings very little, and all three variance percentages remained large and significant (attractiveness = 26.0%, masculinity = 31.4%, dominance = 30.8%). This small change (0.8% on average across the three traits) might be due to the fact that we tested a narrow range of focal

person ages (all 18-24 years old) in the current study; thus, all of the partners tended to be young adults.

Finally, although eliminating the non-committed partners from the dataset reduced the number of usable focal persons (i.e., focal persons with two or more partners) to $N = 50$ ($N = 142$ partners), conclusions remain unchanged in this alternative analysis. All three variance percentages remained large and significant (attractiveness = 44.0%, masculinity = 29.8%, dominance = 44.9%). In other words, clustering of attractiveness, masculinity, and dominance by focal person is substantial whether we examined all of the partners with whom each focal person has had a romantic or sexual connection or just the partners with whom each focal person formed a committed relationship.

Discussion

Study 1 revealed substantial clustering on traits observable in photographs: A focal person's romantic partners tended to exhibit similar levels of attractiveness, masculinity, and dominance to one another. Some of this clustering was reflected in assortative mating, as the focal person's qualities tended to be associated with their partners' qualities. Furthermore, clustering did not seem to be affected by the inclusion of noncommitted romantic partners in the analyses; past partners exhibited similarities regardless of whether or not their relationship with the focal person was casual or serious. This latter finding is consistent with perspectives noting that, during the partner selection process, people may have difficulty differentiating between partners that prove to be casual and short-term versus committed and long-term (Eastwick et al., 2016).

This study suggests that stable, predictable forces differentiate some focal persons' past and current partners from other focal persons' partners. Mate value is a strong candidate in

explaining the current findings: People reach strong consensus about desirable traits like attractiveness when initially getting to know each other, and sorting on attractiveness tends to emerge if people form relationships during this romantically competitive early phase of the acquaintance process (Eastwick & Buck, 2014; Eastwick & Hunt, 2014; Hunt et al., 2015). Thus, it is a reasonable bet that the high mate value focal persons in this study had multiple experiences attracting partners who also had high mate value in impression formation contexts (e.g., when arriving at college). This explanation could apply to all three observable qualities in the present report (i.e., attractiveness, masculinity, and dominance), which all tend to be associated with romantic desirability and have been studied extensively in the evolutionary literature on mating (e.g., Little et al., 2011; Maner et al., 2008).[3] Although age did not account for the clustering we observed in this study, our participants were from a narrow age range; thus, it is unknown the extent to which these findings (especially the age covariate analysis) would generalize beyond this sample.

Study 2

Study 2 draws from Add Health, a study that tracked thousands of participants across four waves of data collection (Harris & Udry, 1994-2008). At three of these waves, the researchers asked participants to nominate their past and current romantic partners. Many of the partners nominated at the first and second time points were Add Health participants themselves, and some of the partners nominated at the third time point were recruited to participate in the study. Therefore, the Add Health dataset can be reorganized such that romantic partners from up to three time points are nested within focal person, much like the organization of Study 1.

Add Health was designed by sociologists, and as is common with many sociological datasets, it is nationally representative: the researchers collected information about young adults

from many different demographic contexts. Study 2 capitalizes on this feature of the data to separate out the effect of active and passive stable forces by subtracting variance accounted for by each participant's school—that is, their local demographic context. Study 2 also contains a mixture of attributes related to mate selection that have been examined across the evolutionary psychological, sociological, and close relationships literatures.

Method

Participants

Focal persons were the 574 Add Health participants (303 men, 271 women) who nominated two or more romantic partners for whom self-report data were available (i.e., all available participants were used). Add Health is a nationally representative study that used a stratified, school-based sampling design. A comprehensive in-home interview of N = 20,744 adolescents (M_{age} = 16.1) was conducted during the 1994-1995 school year (Wave I) and continued for three additional waves of data collection during 1995-1996 (Wave II), 2001-2002 (Wave III), and 2007-2009 (Wave IV). The focal persons analyzed in the current study all nominated two or more romantic partners across Waves I-III. (The Add Health research team did not assess partner data at Wave IV.)

Approximately half of the Add Health participants (N = 11,352) nominated two or more partners across Waves I-III. However, for the partners to be usable in the current set of analyses, the partners needed to have provided self-report data, and the majority of the nominated partners were not Add Health participants and provided no such data. Thus, the usable data consisted of the 574 focal persons nominating two or more romantic partners who *also* happened to provide self-report data as part of the study. The 574 focal persons attended 95 different schools at Wave I. In terms of race/ethnicity, 70.0% identified as White, 17.2% as Black or African-American,

12.4% as Hispanic or Latino, 7.0% as Asian or Pacific Islander, 3.7% as American Indian or Native American, and 8.0% as Other. (Participants could select multiple responses to the race/ethnicity items.) Although the racial composition of this subsample is similar to the overall Add Health nationally representative sample, the 574 individuals who happen to be usable in the current analyses cannot be considered a random subset of the original Add Health cohort.

The 574 focal persons were romantically involved with a total of 1,110 different partners over the course of the study, some of whom were nominated more than once (i.e., by two or more focal persons). Most focal persons (466) nominated two partners, 93 nominated three partners, 8 nominated four partners, 6 nominated five partners, and 1 nominated six partners, which resulted in a dataset of 1,279 partner-reports. In the analyses reported below, these partner-reports are nested within focal person, which (in some analyses) are subsequently nested within school.

In Study 1, we removed the two focal persons who had same-sex relationships, as their inclusion substantially affected the findings. In this study, we again tested whether the inclusion of data from the 12 focal persons who reported same-sex relationships substantially affected the findings; the percentage values in Table 2 changed a mere 0.4% on average when these individuals' same-sex reports were excluded. Thus, all eligible focal persons were included in the dataset regardless of sexual orientation.

Procedure and Materials

Romantic partner nominations. At Waves I and II, Add Health participants nominated up to three individuals with whom they had a "special romantic relationship" over the previous 18 months and up to three additional individuals with whom they had been sexually involved (i.e., six possible total nominations at both time points). At Wave III, a subset of the Add Health

participants were asked to recruit their current romantic partners to join the Partner Sample, which ultimately consisted of one-third married, one-third cohabiting, and one-third dating couples. Of the 1,279 partner-reports used in the analyses below, 625 were Wave I nominations, 445 were Wave II nominations, and 209 were Wave III nominations.

Partner individual difference variables. We identified seven measures in the Add Health dataset that are broadly relevant to partners' romantic desirability according to evolutionary psychological or close relationships perspectives (educational aspirations, depression, intelligence, self-esteem, vitality; Buss & Barnes, 1986; Brase & Guy, 2004; Fletcher et al., 1999; Kirsner et al., 2003) and/or linked social homogamy perspectives on sexual activity and romantic relationship formation in emerging adulthood (delinquency, depression, religiosity; Harden & Mendle, 2011; Joyner & Udry, 2000; Martin et al., 1986; Taylor, McGue, & Iacono, 2000; see Supplementary Table 1).[4] Each construct was standardized within wave for analyses. Add Health item codes are included in the Supplementary Material.

Six of the seven constructs were assessed with self-report scales. Partners indicated their *delinquency* using a 15-item (Wave I), 14-item (Wave II), or 13-item (Wave III) log-transformed measure (e.g., "In the past 12 months, how often did you steal something worth more than $50?"; $\alpha = .83$); *depression* in the past week using a 19-item (Waves I and II) or 9-item (Wave III) measure (e.g., "You felt depressed"; $\alpha = .86$); *educational aspirations* (Waves I and II only) using a 2-item measure ("How much do you want to go to college," "How likely is it that you will go to college?"; $\alpha = .82$); *religiosity* using a 3-item measure (e.g., "How important is religion to you"; $\alpha = .90$); *self-esteem* using a 6-item (Waves I and II) or 4-item (Wave III) measure (e.g., "You have a lot to be proud of"; $\alpha = .85$); and *vitality* (Waves I and II only) using a 2-item measure ("I have a lot of energy," "I am physically fit"; $\alpha = .64$).

The seventh construct was an *intelligence* (*IQ)* test that partners completed at Wave I. This test (the Adolescent Health Picture Vocabulary Test) is a 78-item abbreviated, computerized version of the Peabody Picture Vocabulary Test, a measure of verbal intelligence. IQ scores were available for all Wave I and II partners.

We also collected data to address the extent to which these variables were normatively versus idiosyncratically desired in a partner. To assess desirability, we asked a separate sample of 201 Mechanical Turk participants (95 men, 102 women, 4 transgender; M_{age} = 33.6) to rate the importance of each of the 7 desirability variables in a mate. The instructions prompted 99 of the participants to consider how important each quality is to other people in general, whereas the remaining 102 participants evaluated how important each quality is to people currently in high school (i.e., the age of the participants in our sample). Items were worded identically to the corresponding Add Health items but beginning with the phrase "Someone who" (e.g., "Someone who feels that he/she has a lot to be proud of"). For self-esteem, depression, and delinquency, we only assessed the three highest loading items on the construct (based on a factor analysis of each construct at Wave I using principal axis factoring and extracting a single factor); ratings for each item were averaged for analyses. Ratings were made on a scale from -9 (*no one thinks this person would be a valuable mate*) to 0 (*some people think this person would be a valuable mate*) to 9 (*everyone thinks this person would be a valuable mate*).

These ratings suggested that five of the seven qualities we assessed were especially normatively desirable: a lack of delinquency, self-esteem, vitality, educational aspirations, and intelligence (Supplementary Table 2). Two attributes—depression and religiosity—were believed to be more desirable to some people than to others.

Partner demographics. We also examined three variables traditionally representative of socioeconomic status. Two were completed by the parents of the partners at Waves I and II only: *Parental income* using a 1-item log-transformed measure ("About how much total income, before taxes did your family receive in 1994? Include your own income, the income of everyone else in your household, and income from welfare benefits, dividends, and all other sources."); and *parental education* using a 2-item measure ("How far did you go in school?" and "How far did your current spouse/ partner go in school?"; $\alpha = .75$). The third was *partner White race*: whether or not the partner indicated that his/her race was White.

Analysis Strategy

As in Study 1, we calculated the focal person ICC for each of the seven individual difference constructs across all partner-reports. In the dataset, each row consisted of a single partner, and an indicator variable (i.e., focal person ID) linked each partner to his/her focal person. In this study, we calculated the ICC using two methods: The first method treated focal person ID as a random factor, and the second method treated both school and focal person ID nested within school as random factors. In both cases, the statistic of interest is the random variance estimate for the focal person ID divided by the total variance. The statistic provided by the first method indicates the percentage of variance in partners' characteristics due to the focal person; that is, to what extent are partners with a particular characteristic more likely to be clustered around some focal persons rather than others? In this study, however, some of this variance could be due to the fact that some focal persons are more likely than others to encounter partners with particular qualities in their daily lives (i.e., social homogamy). To account for this possibility, the second method subtracts the percentage of variance accounted for by the focal person's school. Thus, the second method provides a measure of the extent to which partners

with the particular characteristic are clustered around focal persons given the existing range of values on that characteristic in the focal person's immediate environment. As in Study 1, all analyses control for the sex of the focal person. Data for Study 2 cannot be shared publicly per restricted use data contract with the University of Michigan.

Results

Main Analyses

Table 2 presents the percentage of variance accounted for by focal person across the seven partner qualities. The first column of data presents these variances without controlling for school; these variances therefore indicate the extent of clustering due to stable active and passive selection forces combined. In these analyses, the percentage of variance due to the focal person ranged from 6.5% (vitality) to nearly 30% (IQ); the average variance across all qualities was 12.9%, and all except vitality were significantly greater than zero. In other words, when active and passive selection forces are taken into account, there was some clustering of partner qualities by focal person: Some people tended to have romantic partners with high levels of particular qualities, whereas other people tended to have romantic partners with low levels of these qualities. Nevertheless, with the exception of IQ, these variances were considerably lower than the variances assessed for the photograph-rated qualities in Study 1.

In a second analysis that isolated the influence of active selection forces, these variances were reduced considerably. Controlling for school reduced the focal person variance (second column) in all cases. The average variance across all seven qualities was 5.2%, and only delinquency and self-esteem were significantly greater than zero. For several of the desirable qualities that are typically ranked highly on lists of ideal partner preference qualities (e.g., intelligence, educational aspirations, vitality; Fletcher et al., 1999), the percentage of variance

due to the focal person was extremely low, and none reached a "meaningful" threshold of 10% (Kenny et al. 2006). In summary, above and beyond the effect of location (i.e., school), there was only weak evidence that some people were more likely than other people to form romantic relationships with partners who possessed these qualities.[5]

We have depicted the effect of school in this study as a passive force. In principle, the effect of school could partially reflect an active force if focal persons were able to persuade their parents to move them to schools that contained potential partners with attributes they desired. If this circumstance were common, then controlling for school inappropriately subtracts active forces in addition to passive forces from the analysis. To address this concern, we examined only the $N = 321$ focal persons who reported at Wave I that they had moved to their current home before age 11 (an age before which it would be nearly inconceivable that children would persuade their parents to move to a new school for the purposes of dating). The same pattern of results emerged in this analysis: The average variance across all qualities was 10.1% without controlling for school and 3.7% after controlling for school, and none of the variances after controlling for school were significantly different from zero. This analysis supports our assumption that the school variable reflects passive forces in this sample.

For comparison purposes, we also examined the percentage of variance in partner demographics attributable to the focal person, with the expectation that the school that the focal person attended should strongly account for clustering among partners on these variables. Without controlling for school, these percentages were moderate to large: Focal person variance was 17.2% for parental income, 17.9% for parental education, and 75.8% for partner White race. In other words, some focal persons were more likely than others to form romantic relationships with partners who possessed particular demographic backgrounds. After controlling for school,

however, the percentages for parental income and parental education were negligible (0.0% and 0.6%, respectively). Even after accounting for school stratification, there was still evidence that people's romantic partners were somewhat similar with regards to race; focal person variance was 14.8% for White race. Adolescents and young adults tend to be romantically and/or sexually involved with partners who are demographically similar to each other, because these partners are drawn from educational niches that are demographically stratified, rather than because of active selection processes.

All seven constructs were also available for the focal persons themselves, and we could therefore examine the extent to which similarity played a role in predicting clustering. Similarity effects are presented in the fourth column of Table 2; these values indicate the standardized beta for the focal person's score on the attribute predicting the partners' scores on the attribute (as a fixed effect) without accounting for the random effect of school (i.e., the original analysis above). Three of the seven similarity effects achieved an effect size that was at least small (i.e., $r = .10$): educational aspirations, religiosity, and IQ. In other words, focal persons tended to form relationships with partners who were similar to them on educational aspirations, religiosity, and IQ. The evidence for similarity with respect to depression, self-esteem, vitality, and delinquency was quite weak, although all seven variables had positive correlations that were significantly different from zero. Not surprisingly, controlling for similarity caused variance due to the focal person to decline in proportion to the size of the similarity effect (e.g., a reduction of 0.2% for delinquency vs. a reduction of 7.9% for IQ).

Intriguingly, the similarity correlations tended to be higher for the attributes that exhibited large proportions of variance at the school level. Although the *N* is only seven qualities, the correlation between the similarity betas and the school variance % in Table 2 is $r =$

.994. Also, controlling for school markedly reduced the three correlations with school variance above 10% (religiosity, educational aspirations, and IQ; see fifth column in Table 2). These analyses tentatively suggest that similarity effects tended to emerge when the attribute in question exhibited demographic stratification; that is, intelligent focal persons are more likely to date intelligent partners because they attended schools with intelligent people, but delinquent focal persons were no more or less likely to date delinquent partners because delinquency does not cluster at the school level.

Subsidiary Analyses

One possible alternative explanation presents itself when considering Table 2. In principle, the Wave III partners could be responsible for the low values observed. Unlike the Wave I and II partners, these partners were not part of the original Add Health cohort, nor did they necessarily attend the same high school as the focal person. To investigate this possibility, Table 3 presents the percentage of variance accounted for by the focal partner for the seven partner qualities with Wave III partners removed from analyses. (Educational aspirations, vitality, and IQ were not assessed for the Wave III partners, so these values remain unchanged from the Table 2 values.) Deleting Wave III partners had few overall effects on the percentages. Analyses conducted without controlling for school suggested a moderate amount of variance due to the focal person (average variance = 14.8%), but this variance dropped when controlling for school (5.7%). A handful of qualities exhibited somewhat higher focal person percentages in these analyses (e.g., religiosity without controlling for school), but on average, the absolute value change in the percentages was small (~5.8% across both columns for delinquency, depression, religiosity, and self-esteem).

As in Study 1, we examined whether our inclusion of the sexual partners had a substantial impact on the results. Eliminating the Wave I and II sexual partners from the dataset reduced the number of usable focal persons (i.e., focal persons with two or more partners) to $N = 446$. Again, the results changed very little (Supplementary Table 3). Across the seven qualities, the values that emerged from this reduced analysis were very similar to the percentages indicated in Table 2. Analyses conducted without controlling for school revealed a small amount of variance due to the focal person (average variance = 11.3%), but this variance dropped by half when controlling for school (5.1%). The average absolute value change in the percentages relative to Table 2 was again small (~2.2% for all seven qualities across both columns). In other words, to the extent that there is a meaningful distinction between a "special romantic relationship" and a sexual partner at Waves I and II, this distinction is unlikely to be responsible for the low values in Table 2.

Waves I-III of the Add Health study spanned eight years, and if the personality, mate preferences, or mate value of focal persons changed over this time period, ICCs would naturally be low even if these factors affected mate selection strongly at the time the partners were selected. Therefore, the stable factors that characterize focal persons might produce ICCs of a substantial effect size if ICCs were calculated over a shorter time period that permitted less change. To address this possibility, we calculated ICCs at Wave 1 ($N = 208$ focal persons, 449 partners) and at Wave 2 ($N = 92$ focal persons, 196 partners) for those focal persons who possessed more than one partner within the wave. (Wave 3 could not be analyzed separately because no focal person contributed more than one partner to Wave 3.) Thus, these analyses were limited to partners whom focal persons dated no more than 18-months apart. Results were largely consistent with those reported above for the whole sample (Supplementary Tables 4 and 5). Overall, analyses conducted without controlling for school revealed a moderate amount of

variance due to the focal person (average variance = 18.6% across Wave I and Wave II); this variance dropped by half when controlling for school (10.3%), although this amount of variance does indicate a small average effect. The average absolute value change in the percentages was modest (~8.0% across both columns for both waves) relative to the whole sample analyses. In short, over a brief span of time on the dating market (i.e., less than a year on average), stable characteristics of focal persons accounted for a small portion of the variance in the qualities of their romantic partners, although these analyses naturally have a smaller *N* than the whole sample analyses reported above and may be less replicable.

One limitation of the ICC approach in all of the studies in this article is that they require that we examine a single attribute at a time. In this study, we also examined an alternative approach in which we calculated a within-focal-person correlation across all seven attributes for each pair of romantic partner nominations. For example, if a focal person contributed two partners to the dataset, we calculated the correlation between the two partners' qualities across all seven variables. This pattern metric (Eastwick & Neff, 2012) assesses the extent to which a focal person's romantic partners exhibited the same pattern of traits across all qualities, regardless of the level of the traits. For focal persons with more than two romantic partner nominations, we averaged the correlations calculated across all possible pairs (i.e., 3 correlations for 3 nominations, 6 correlations for 4 nominations, etc.). Across all focal persons, this within-person correlation was small: *Mean r* = .12, *Median r* = .14, *SD* = .48. Conclusions about these values should remain tentative given that (a) the extent to which they would decline when controlling for school remains unknown and (b) these within-focal-person correlations were calculated on only seven attributes. Nevertheless, we present these values here in the interest of

completeness and as an example of how researchers could take a multi-attribute approach to this question.

Discussion

The Study 2 Add Health data did reveal some clustering with respect to partner qualities; although the average effect size was small (~13%), IQ was particularly large (nearly 30%). In other words, focal persons' past and current romantic partners tended to exhibit some similarities to one another, just as in Study 1. But after accounting for demographic stratification, the extent of clustering dropped to 5% on average (i.e., a trivial effect size). In other words, clustering on the Add Health attributes emerged because people are dispersed across environments that are demographically stratified, not because their stable qualities (e.g., mate preferences or mate value) aid them in selecting partners who possess different levels of these attributes within their local pool of available mates. Conclusions did not vary substantially if we examined romantic partners alone (rather than "romantic" and "sexual" partners combined) or if we examined narrower time frames that might have accentuated the effects of stable forces. Even for demographic variables, the variance in partner choice that could be explained due to active factors ranged from small (partner White race) to near zero (parental income and education).

Rather, clustering in this study was primarily due to passive partner selection forces; focal persons and their partners tend to be sorted into different schools, and focal persons seemed to select partners unpredictably within this local pool with respect to the romantically relevant qualities contained in the Add Health dataset. Consistent with this logic, assortative mating tended to emerge for those qualities that exhibited clustering at the level of the school: When schools differed on a particular variable (e.g., IQ), focal persons and their partners exhibited

similarities on that variable, and when schools did not differ on a variable (e.g., delinquency), focal persons and their partners did not exhibit similarities.

In summary, sociologists would conclude from the Study 2 data that much of the partner selection process is driven by demographic stratification, and they would be correct—given the variables that they tend to investigate in datasets like Add Health. Of course, these conclusions might shift among older individuals who have greater control over their living context (e.g., people who have the means and freedom to move to areas where they might encounter partners with qualities they desire; Motyl, Iyer, Oishi, Trawalter, & Nosek, 2014). But the participants in this study were largely older adolescents and young adults who likely had little choice over their living context. Furthermore, given that the Add Health researchers collected data from a representative sample of U.S. schools, it is plausible that many of the parents of the participants in this study would have faced practical and financial challenges moving to new neighborhoods even if they had desired to do so (Orr et al., 2003). Thus, we posit that the clustering observed here is largely passive; focal persons dated similarly intelligent partners because they happened to encounter partners of similar levels of intelligence.

Study 3

In some cases, clustering among a focal person's past and present partners tells us something about the partner selection process: Were focal persons acting on their mate preferences, leveraging their mate value to obtain desirable mates, and/or constrained by their local demographic contexts? But examining clustering among a focal person's past and present partners can also reveal something about what takes place as relationships form and develop—the process by which two individuals assess how they feel about a particular partner and what they want from that relationship (Huston, Surra, Fitzgerald, & Cate, 1981; Knapp, Vangelisti, &

Caughlin, 2014). Importantly, this relationship formation process is typically gradual: Normative trajectories of romantic interest rise over time as potential partners engage in sexual behaviors and assess their emotional and physical chemistry (Finkel, Simpson, & Eastwick, in press). Indeed, the point at which a relationship starts to become sexual (e.g., the first "make out") is approximately the point at which people begin to lose interest in short-term flings and experience increasing interest in partners that have long-term potential (Eastwick et al., 2016). In other words, dating, hooking up, becoming exclusive, and building commitment reflect different components of a gradual evaluative process in which potential romantic partners try to gauge how positively they feel about each other; some of these relationships progress to later stages, and others do not.

Stable forces could also operate during this process in the form of *relationship aptitude* (Finkel et al., 2012; Karney & Bradbury, 1995), which refers to stable, enduring strengths or vulnerabilities that persist across different relationships and cause some people to be evaluated more positively than others as relationship partners. The relationship aptitude construct is similar to the concept of mate value, but it is typically operationalized not as the possession of romantically desirable traits but rather the ability to inspire positive relationship outcomes (e.g., the extent to which a focal person is judged by partners to be a desirable/satisfying romantic partner). Partners do not necessarily need to be in a committed relationship with a focal person in order to make judgments relevant to relationship aptitude; they merely need to be able to evaluate how positively they feel about being a romantic relationship with him/her (Eastwick & Hunt, 2014).

The lack of data deriving from a given focal person's multiple romantic partners over time has hindered scholars' ability to assess the extent to which relationship aptitude operates as

a stable force on relationship outcomes. Imagine a focal person who dates several different partners over a period of a few years. On the one hand, any enduring characteristic of the focal person that caused one of his romantic partners to have a poor experience with him should presumably carry forward and cause future romantic partners to rate him similarly poorly (e.g., a neurotic or anxiously attached focal person should have partners who find him unsatisfying as a partner; Campbell, Simpson, Boldry, & Kashy, 2005; Donnellan, Conger, & Bryant, 2004; McNulty, 2013). On the other hand, relationship-specific judgments are quite idiosyncratic. For example, one study that asked opposite-sex friends and acquaintances to rate each other on such measures found that consensus was essentially zero (Eastwick & Hunt, 2014); that is, as people get to know each other better over time, people exhibit very little agreement about who is romantically desirable and who is not. Therefore, to the extent that one partner's evaluation of a focal person is driven by subjective, affectively-based factors that vary considerably from partner to partner, relationship aptitude might not be sufficiently stable to produce clustering in the relationship-specific judgments (e.g., satisfaction, romantic desirability) of one's actual past and present romantic partners.

A proper examination of this question requires data from multiple romantic partners, just as in Studies 1 and 2. But instead of examining those partners' traits, this question requires partners' relationship-specific judgments about a common focal person. This type of variable is common in the literature on relationship initiation (Sprecher & Duck, 1994) and on established close relationships (Fletcher et al., 2000a, 2000b), where researchers frequently investigate subjective reports provided by one partner about the other partner (or the relationship). Until now, no dataset that includes such measures reported by multiple partners about the same focal person has been published.

In order to address the question of whether relationship aptitude is a stable factor that affects people's relationship experiences over time, we took advantage of a unique database that allowed us to assess the extent of clustering in target-specific reports. In 2014, a new website harnessed the pervasive tendency among young adults to publicly rate things, places, and people by creating an online forum for women to rate their past and present romantic partners on a number of romantic dimensions. The website was a romantic ratings database unlike any that had ever existed because these ratings consisted of women's judgments of men *whom they actually knew*. (No online dating website can make this claim, for example, because evaluations on dating websites are based only on profiles, not live interactions; Finkel et al., 2012.) The website later ran into legal complications and today is not as fully open as it once was. But before the website changed its format, we downloaded ratings data on 400 focal persons to examine several different research questions, including the following: Do romantic desirability judgments exhibit clustering when women evaluate men with whom they had actually had a sexual experience?

Method

Participants

These data were taken from a study of 400 men who were rated on a website designed for women to share their honest opinions about different men they knew personally. Of these 400 focal persons, 145 (M_{age} = 23.0 years, *SD* = 3.9, range = 18-36) happened to be rated by two or more women who identified themselves as current/former romantic partners (see below); analyses below were conducted on this subsample.

At the time that the data were collected by the research team (early-to-mid 2014), the website was publicly accessible; visitors to the website viewed Facebook profile photos of men alongside evaluation forms of the men completed by various women. The men did not have to

provide consent to be included on the website; they merely had to possess a Facebook account. The identities of the women who rated the men were completely anonymous.

A research assistant browsed the website and included a man in the sample of focal persons if he met all of the following criteria: (a) The photograph clearly depicted only one man (so that the focal person could be identified clearly); (b) the photograph was clear and not blurry; (c) the shoulders and face of the man were clearly visible; and (d) at least two women provided ratings of the man (otherwise ICCs could not be calculated). Every time that the research assistant encountered a man who passed all four criteria, the assistant would assign him a focal person ID number, download his (publicly available) photograph, record his college, age, and relationship status (presumably gleaned by the website from Facebook), and copy the evaluation forms provided by all the women who rated him into a spreadsheet.

The research assistant evaluated each man for inclusion in the dataset in the order that he appeared when scrolling through the website; this order seemed to be random. (We reached out to the company to verify this hypothesis but received no response.) If no search terms were entered, the men all appeared to originate from the same city as the current user (and many happened to share the same university affiliation, which was the University of Texas at Austin). After approximately 200 focal persons were downloaded, the website no longer presented any new men from Austin, Texas. At this point, the research assistant entered other large, geographically diverse major universities as search terms (Arizona State University, Penn State University, University of Minnesota, University of Florida, Ohio State University, Texas A&M University, and the University of Michigan) until reaching the target sample size of 400 focal persons. The search would often produce men who attended nearby universities and high schools; in the end, the $N = 145$ men attended 40 different schools.

Procedure and Materials

Female website users had the opportunity to complete an evaluation form about any man whom they knew personally (as long as he had a Facebook profile). First, the woman indicated how she knew the man by selecting from the following six options: Friend, crush, hooked up, relative, ex-boyfriend, and together. Only women who selected the hooked up, ex-boyfriend, and together options are included in the analyses reported in this manuscript (N = 429 women rating N = 145 focal men), as these are the women who had some amount of romantic/sexual experience with the man and most closely parallel the samples reported in Studies 1 and 2. (Analyses conducted on the full sample including friends, crushes, and relatives revealed identical conclusions, similar to the findings of Eastwick & Hunt, 2014.)

Second, all women rated the man on five attributes using a 1- to 5-star rating scale: *Appearance*, *humor*, *manners*, *ambition*, and *commitment*.[6] Despite the fact that these items have distinct face validity, a factor analysis (principal axis factoring with promax rotation) and inspection of the scree plot revealed a one-factor solution (explaining 59.0% of the variance), and the alpha on this five-item *romantic desirability* construct was high (α = .82). We use an average of these five items in the analyses reported below, although a construct computed from factor loadings revealed identical conclusions. Women who selected the categories hooked up, ex-boyfriend, or together (i.e., all the women in the current set of analyses) also completed a two-item *sexual satisfaction* measure using the same 1- to 5-star rating scale consisting of the items *kissing* and *sex* (α = .78).

Third and finally, the women selected his *best qualities* (e.g., #NoIssuesHere, #Trustworthy, #Trailblazer, #CaptainFun, #CuddlesAfter) and *worst qualities* (e.g., #NeverLetsMeWin, #DeathBreath, #HeLovesMeNot, #StripClubVIP, #NoStyle) from a set of

several hundred predefined hashtags. The research assistant added up the number of positive hashtags (M = 11.4, SD = 13.1) and negative hashtags (M = 5.5, SD = 6.5) selected by each female rater; the content of the hashtags that the woman selected was not recorded.

Analysis Strategy

As in Studies 1 and 2, we calculated the focal person ICC for each construct across all reports made by women who identified their relationship with the man as hooked up, ex-boyfriend, or together. In the dataset, each row consisted of a single partner, and an indicator variable (i.e., focal person ID) linked each partner to his/her focal person. All the focal persons were men in this study, so there was no need to control for focal person sex. Data and code for Study 3 are available here.

Results

Main Analyses

Table 4 presents the percentage of variance accounted for by focal person for the overall romantic desirability construct, the sexual satisfaction construct, the number of best qualities selected, and the number of worst qualities selected. The percentage of variance due to the focal person was small, ranging from 3.2% (number of negative qualities) to 8.6% (romantic desirability); the average variance across all four constructs was 6.2%, and none was significantly greater than zero. In other words, there was little or no clustering by focal person: Female partners tended not to agree whether the men with whom they had romantic/sexual experience were romantically desirable, were sexually satisfying, or how many positive and negative qualities they had. Rather, these sorts of ratings seem more likely to be a function of rater qualities (i.e., some women have negative feelings about their ex-boyfriends whereas others

do not) and relationship qualities (i.e., the men were romantically appealing to some women but not to others; Eastwick & Hunt, 2014).

The single item "appearance" parallels the Study 1 attractiveness construct, and so it is potentially valuable to examine clustering on this item alone. In contrast to Study 1 (which found 27% clustering in the attractiveness of current and ex romantic partners), the appearance judgments made by current and ex romantic partners in the current study revealed clustering of only 6.5%, a value that was not significantly different from zero. That is, a man's current and ex romantic partners might be similarly attractive according to independent coders (i.e., Study 1), but those current and ex romantic partners themselves do not agree whether the man himself is attractive or unattractive (i.e., Study 3).

In contrast to the Study 2 data, adding school as an added layer of nesting did nothing to change the findings: For all four analyses, school accounted for 0.0% of the variance, and the values in Table 4 remained identical. In other words, men from some schools were no more or less likely to elicit positive ratings on these measures than men from other schools.

Subsidiary Analyses

Some of the ratings in the current dataset were provided by women about their *current* male partners. Given the well-known tendency for people to rate their romantic partners extremely highly on evaluative qualities (Fletcher & Kerr, 2010), it is possible that these positively biased women contributed to the low variabilities in Table 4. Indeed, the $N = 34$ women who selected "together" significantly differed from the remaining women (who selected "hooked up" and "ex-boyfriend") on romantic desirability, $t(427) = 5.37$, $p < .001$, $d = .52$, sexual satisfaction, $t(417) = 4.35$, $p < .001$, $d = .43$, number of best qualities, $t(427) = 3.85$, $p < .001$, $d = .37$, and number of worst qualities, $t(427) = -3.06$, $p = .002$, $d = -.30$. Therefore, we

reconducted the analyses subtracting the romantic partners from the dataset (for a similar analysis, see Eastwick & Hunt, 2014). Eliminating the committed partners from the dataset reduced the number of usable focal persons to $N = 130$ ($N = 381$ partners). Conclusions remain unchanged in this alternative analysis, as all four variance percentages remained small and nonsignificant (romantic desirability = 8.2%, sexual satisfaction = 2.3%, best qualities = 4.9%, worst qualities = 1.2%). In other words, the focal men's ex-girlfriends and hookup partners did not agree whether each man was romantically desirable, was sexual satisfying, or whether he had many positive and negative qualities.[7]

The Social Relations Model (Kenny, 1994) notes how variables like those contained in the current dataset (i.e., a rater rates a specific other person) consist of four sources of variance: actor variance (i.e., the extent to which raters differ in how they rate targets on average), partner variance (i.e., the extent to which raters agree in their ratings of specific targets on average), relationship variance (i.e., the extent to which raters differ in their ratings of specific targets above and beyond actor and partner variance) and error variance (i.e., the extent to which raters differ in their ratings of specific targets due to chance responding). The data in Study 3 are a dyadic design called a nonreciprocal one-with-many design (Kenny et al., 2006). In such a design, it is possible to divide the variance into three sources: partner variance (which is conceptually identical to the ICCs calculated above), error variance, and a variance estimate that includes the sum of actor and relationship variance. (A nonreciprocal design cannot further separate actor and relationship variance, although a reciprocal design can do so, see Eastwick & Hunt, 2014.)

We wished to obtain a cleaner estimate of the extent of clustering due to the focal person (i.e., partner variance) using a Social Relations Model design that separated partner variance,

error variance, and actor + relationship variance. We adapted the procedure of Eastwick and Hunt (2014, Study 3) that separated each multi-item construct (i.e., romantic desirability, sexual satisfaction) into two "bins" on separate rows of the dataset (e.g., Betty's ratings of Dane's appearance, humor, and manners were averaged to form one bin and her ratings of Dane's ambition and commitment were averaged to form the second bin). Then, we ran a multilevel model that produced variance estimates for the focal person (i.e., partner variance), the focal person × rater interaction (i.e., actor plus relationship variance), and error. For romantic desirability, partner variance was 7.0%, actor plus relationship variance was 59.1%, and error was 32.9%.[8] For sexual satisfaction, partner variance was 5.1%, actor plus relationship variance was 57.5%, and error was 37.5%. In short, this cleaner estimate of partner clustering (i.e., partner variance calculated after subtracting error variance) was nearly identical to the ICCs reported above.

Finally, as described above, the N = 145 focal persons examined here come from a larger sample of N = 400 focal persons. Some of the focal persons in this larger sample (N = 257, M_{age} = 22.6) received reports from at least two female partners who classified themselves as friends. These friends completed the same romantic desirability, positive qualities, and negative qualities measures (but not the sexual satisfaction measure) as the current and ex romantic partners described above. In principle, we can calculate clustering on these reports (N = 832 total reports) as well. In these analyses, the percentage of variance due to the focal person was again small for romantic desirability (7.3%); this value provides a conceptual replication of Eastwick and Hunt (2014, Study 3), which found similar values (0.0-7.5%) for target variances reported by opposite-sex friends and acquaintances across several similar measures. Clustering for the number of negative qualities (8.2%) but was also low but was (unexpectedly) medium-sized for number of

positive qualities (19.5%). Opposite-sex friends may reach agreement about each other's positive qualities more easily than they reach agreement about negative qualities or romantic desirability.[9]

Discussion

The results of Study 3 revealed very little evidence for clustering when multiple past and present romantic partners completed target-specific reports about a focal person. That is, a man's exes and current partners did not agree whether he was romantically desirable, whether he was a good sex partner, or whether he possessed many positive and negative qualities. The measures in this study did not exhibit any clustering at the level of school; the Study 2 Add Health data might have found clustering at the school level because those schools exhibited great demographic diversity, or it is possible that the school that one attends is related to variables like IQ and educational aspirations but not the ability to inspire romantic desire and other variables commonly investigated by close relationships researchers. Finally, the level of clustering remained very low even when current romantic partners were removed from the sample, suggesting that the positive biases exhibited by current romantic partners were not the (sole) cause of these low percentages.

Stable individual differences that imbue some focal persons with more relationship aptitude than others may not persist in affecting the evaluations of their multiple romantic relationships over time. On the surface, this implication may seem shocking in the face of the myriad effects of personality on people's experiences in relationships (McNulty, 2013). But bear in mind that most of this prior literature has documented *rater* effects (e.g., a person's neuroticism negatively predicts his own satisfaction; Robins et al., 2002), and only *target* effects (e.g., a person's neuroticism negatively predicts his partner's satisfaction) are directly relevant to

the clustering assessed in this study. Furthermore, even if the current results seem surprising, low consensus on affective measures has strong precedent in the prior literature. Among long-term acquaintances, highly affect-laden measures (e.g., supportiveness) exhibit extremely low levels of consensus (e.g., Lakey, McCabe, Fisicaro, & Drew, 1996) even though consensus on personality judgments is typically strong in these contexts (Kenny, Albright, Malloy, & Kashy, 1994). In addition, Eastwick and Hunt (2014) documented low levels of consensus in two studies when opposite-sex friends and acquaintances rated each other with respect to romantically desirable traits and "estimates" of how satisfied they would be in a romantic relationship with each other. The current study suggests that current and past romantic partners—people who actually have romantic experience with the focal person—seem to achieve similarly low levels of agreement about whether a particular focal person is a desirable partner or not. The subjective, dyad-focused reports that are commonly used in the close relationships literature may be highly idiosyncratic and not strongly influenced by stable, predictable features of the person being evaluated.

General Discussion

In principle, heterosexual individuals in most contemporary societies could form romantic relationships with a vast number of peers. But they will only ever meet a subset of those peers—a subset that historically has been circumscribed by a demographically specific local context. Furthermore, people experience the desire to become romantically involved with only some of the opposite-sex individuals whom they know, and only a portion of this select group will reciprocate that desire. In combination, these elements whittle down each person's universe of possible pairings to a unique pool of current and ex-romantic partners.

Consistencies and inconsistencies in the observable attributes (Study 1), self-reports (Study 2), and target-specific ratings (Study 3) provided by a focal person's unique pool of partners may help scholars to achieve novel insights into this whittling process. Substantial consistency emerged in Study 1 with respect to observable attributes of romantic partners: A focal person's partners were similarly attractive, masculine, and confident, and all three of these qualities exhibited assortative mating (i.e., focal persons' attributes correlated with their partners' attributes). Consistency also emerged in Study 2 with respect to some self-reported qualities of romantic partners (e.g., IQ, educational aspirations). But when consistency did emerge, it was largely due to demographic sorting into different school contexts, and when assortative mating emerged, it also appeared to be driven by this passive mate selection process. Finally, Study 3 revealed little evidence for consistency in the target-specific ratings provided by partners about a focal person: Past and present partners do not agree about a focal person's desirability, sexual satisfactoriness, or his positive and negative qualities. Furthermore, school context could not account for variance in these subjective ratings by partners about focal persons.

How can these findings aid in researchers' attempts to stitch together three disparate literatures on mate selection? The following depiction fits cleanly with the current results and with the prior literature: As sociologists have long emphasized (Kalmijn, 1998, Schwartz & Mare, 2012), "mating requires meeting" (Kalmijn & Flap, 2001, p. 1289), and therefore people are likely to encounter potential romantic partners who are similar to them on attributes that vary by living context. Some of these attributes are normatively desirable (e.g., IQ, educational aspirations), where other attributes are desirable to some people but not to others (e.g., religiosity), but people will be more likely to meet (and ultimately form relationships with) partners who are similar to them on attributes that vary by location. Within these

demographically similar pools, potential romantic partners get to know each other; initially, there is high consensus about which potential mates do and do not have the desirable, observable qualities oft-examined by evolutionary psychologists (e.g., attractiveness, masculinity, confidence; Little et al., 2011; Maner et al., 2008; Rhodes, 2006). Among the relationships that form early in the acquaintance process, the mates who have consensually desirable and observable qualities—the individuals with high mate value—are able to attract partners with consensually desirable qualities (Back et al., 2011; Hunt et al., 2015). But consensually desirable qualities have weaker effects as potential partners get to know each other better (Eastwick & Hunt, 2014), and once romantic relationships actually form, the dyadic processes and motivated biases emphasized by close relationships scholars come to the fore (Murray et al., 1996; Rusbult et al., 2001). As relationships evolve and change for the better or for the worse (Eastwick et al., 2016; Karney & Bradbury, 1995), partners' evaluations of each other are affected less by the stable qualities of the partner being evaluated and more by idiosyncratic, relationship-specific factors.

Process depictions like the one offered here will prove valuable as researchers across evolutionary psychology, sociology, and close relationships work to integrate their studies of mate selection. Traditionally, these three fields have studied mate selection in relative isolation (Durante, Eastwick, Finkel, Gangestad, & Simpson, 2016; Eastwick, 2016), and each has tended to adopt its own methods and measures. The three studies in this article used a single analysis strategy (i.e., calculating ICCs for mating-relevant constructs with respect to current and ex partners) and yet revealed three different conclusions because they examined three different kinds of constructs. Like the three proverbial blind men describing different parts of the elephant, the fields of sociology, evolutionary psychology, and close relationships may have been

describing different parts of the mate selection process. Sociologists describe the forces that determine who enters a person's pool of potential partners but little about what takes place within that pool. Evolutionary psychologists examine what happens when people initially form impressions of each other but less commonly examine how actual relationships develop (Eastwick et al., 2016). Close relationships researchers study how existing couples navigate their relationships over time but rarely examine the events and processes that precede couple formation (Campbell & Stanton, 2014). Future approaches to the study of mate selection should endeavor to link these three lines of inquiry, and the current article demonstrates how a single analysis strategy—estimating the degree of consistency among romantic partners—can apply to each of these perspectives.

Strengths, Weaknesses, and Implications

Over twenty years ago, Lykken and Tellegen (1993) raised a question about the extent to which human mate selection is governed by stable internal forces that characterize individuals versus adventitious forces that are dyad-specific and difficult to predict ahead of time. Yet no prior study had attempted to assess the extent to which partner choices are consistent across time, a paradigm that offers perhaps the most direct assessment of the cumulative effect of all stable active and passive selection processes. One strength of the current study is that we pioneered a method that could address this question, and we showed how it could be used to calculate consistency across a variety of different constructs.

Another important strength of the current article is that all three studies captured focal persons' romantic lives over an extended period of time. This longitudinal feature is critical because *Homo sapiens* is a serial pair-bonding species (Fisher, 1989), and in modern Western contexts, many people will form several relationships of varying lengths throughout their young

adult years. Thus, our method was well suited to exploring how stable qualities affect a person's choices and experiences in romantic relationships across this timespan. But of course, many people may treat their young adult years as a time for romantic experimentation, which could result in highly inconsistent romantic partner choices and thus reduced clustering during this period. In addition, personalities can change as adolescents transition to adulthood (Bleidorn, 2015); a focal person who becomes less neurotic over time might end up with romantic partners who have widely varying opinions of his romantic desirability depending on how neurotic he was when the relationship occurred. One weakness of the present article, therefore, is that these methods cannot test whether stable attributes affect romantic processes over short stretches of time. We were able to examine 18-month timespans in the subsidiary analyses in Study 2, but even this period does not preclude stable attributes from changing. The best test of the possibility that stable attributes affect romantic processes over short time spans would perhaps involve assessing low investment events that do not require the formation of a romantic relationship (e.g., does clustering emerge among the partners who agree to a date with a focal person?). Future research should examine this possibility.

A second weakness of the current article is that, across studies, we did not use all three types of assessment strategies (i.e., coder-ratings, self-reports, and target-specific reports) to examine all 14 constructs. That is, we did not examine coder-ratings of religiosity, only self-reports; we did not examine self-reports of masculinity, only coder-ratings. Our approach essentially combined assessment strategy with construct to reflect the common methodological conventions of the three literatures that we were examining. But the consequence of this decision is that we do not know the extent to which the assessment strategy or the constructs of each literature produced the differing results across studies. One exception was physical

attractiveness, which we examined as a coder-related variable in Study 1 but as a target-specific report in a subsidiary analysis in Study 3. The widely differing results (27% in Study 1 vs. 7% in Study 3) suggest that assessment strategy was primarily responsible for the difference between the findings of Study 1 and Study 3, but this evidence is only suggestive. One option for future research would be to cross the three assessment strategies with construct (e.g., collect coder-ratings of partners' depression, partners' self-reports of depression, and partners' reports of a common focal person's depression) in a single population of focal persons and partners, regardless of whether such ratings reflect any sort of typical methodological convention.

Conclusion

Three literatures—evolutionary psychology, sociology, and close relationships—all devote considerable resources to the study of romantic partner selection and maintenance processes. In some cases, these perspectives make diverging predictions regarding mating processes, and these predictions can be pitted against each other to sharpen theories of relationship initiation, maintenance, and dissolution (e.g., Eastwick, Luchies, Finkel, & Hunt, 2014b; Schmitt 2014). But in many ways, the three literatures simply fail to connect, perhaps because they emphasize different constructs and different methodological conventions. The current article highlights how assessments of the extent of clustering among romantic partners may be one analysis strategy that begins to build toward some integration of these literatures.

Regardless of whether stable forces exhibit strong or weak influences on partner selection and maintenance, important theory-driven research will continue to examine mate preferences, mate value, relationship aptitude, and other stable influences. Explaining the remaining, adventitious proportion of the mate selection variance may require that researchers broaden their theoretical horizons. For example, a possibility deriving from chaos theory is that human mate

selection is driven by nonlinear dynamic processes (Weigel & Murray, 2000), not unlike those that cause some natural phenomena to be extremely difficult (e.g., weather) or impossible (e.g., earthquakes) to predict, even with perfect knowledge of initial conditions. New models will need to grapple with such a partner selection process if it is indeed reflective of a large component of human mating (Joel, Eastwick, & Finkel, 2016). And for a complete theory of mating to cohere, these new models will also need to incorporate the variables and processes highlighted by evolutionary psychological, sociological, and close relationships perspectives.

Author Note

This research was supported in part by National Science Foundation Grant BCS-#1243323 awarded to Paul W. Eastwick. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. We wish to thank the following research assistants: Norin Ali, Christine Edomwande, Raynor Fisher, Nathan Gibson, Kevin Ly, Tania Nguyen, Christine Nott, Alex Powell, Eden Stone, Josephina Thibodeaux-Guillory, and Lisa Truong. This research uses data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill and funded by Grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is due to Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health website (http://www.cpc.unc.edu/addhealth). No direct support was received from Grant P01-HD31921 for this analysis.

Contact information:

Paul Eastwick: eastwick@ucdavis.edu
Paige Harden: paige.harden@gmail.com
Jennifer Shukusky: jshukusky@utexas.edu
Taylor Anne Morgan: tayloramorgan@gmail.com
Samantha Joel: samantha.joel@gmail.com

References

Armour, S., & Haynie, D. L. (2007). Adolescent sexual debut and later delinquency. *Journal of youth and adolescence, 36*, 141-152.

Asch, S. E. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology, 41*, 258-290.

Back, M. D., Penke, L., Schmukle, S. C., Sachse, K., Borkenau, P., & Asendorpf, J. B. (2011). Why mate choices are not as reciprocal as we assume: The role of personality, flirting and physical attractiveness. *European Journal of Personality, 25*, 120-132.

Bailey, D. H., Durante, K. M., & Geary, D. C. (2011). Men's perception of women's attractiveness is calibrated to relative mate value and dominance of the women's partner. *Evolution and Human Behavior, 32*, 138-146.

Bale, C., & Archer, J. (2013). Self-perceived attractiveness, romantic desirability and self-esteem: A mating sociometer perspective. *Evolutionary Psychology, 11*, 68-84.

Belot, M., & Francesconi, M. (2013). Dating preferences and meeting opportunities in mate choice decisions. *Journal of Human Resources, 48*, 474-508.

Belsky, J., Steinberg, L., Houts, R. M., & Halpern-Felsher, B. L. (2010). The development of reproductive strategy in females: Early maternal harshness→ earlier menarche→ increased sexual risk taking. *Developmental psychology, 46*, 120-128.

Bleidorn, W. (2015). What accounts for personality maturation in early adulthood? *Current Directions in Psychological Science, 24*, 245-252.

Bouchard, T. J., Lykken, D. T., McGue, M., Segal, N. L., & Tellegen, A. (1990). Sources of human psychological differences: The minnesota study of twins reared apart. *Science, 250*, 223-228.

Brase, G. L., & Guy, E. C. (2004). The demographics of mate value and self-esteem. *Personality and Individual Differences, 36*, 471-484.

Burgess, E. W., & Wallin, P. (1953). *Engagement and marriage*: Lippincott Philadelphia.

Burriss, R. P., Welling, L. L. M., & Puts, D. A. (2011a). Mate-preference drives mate-choice: Men's self-rated masculinity predicts their female partner's preference for masculinity. *Personality and Individual Differences, 51*, 1023-1027.

Burriss, R. P., Welling, L. L. M., & Puts, D. A. (2011b). Men's attractiveness predicts their preference for female facial femininity when judging for short-term, but not long-term, partners. *Personality and Individual Differences, 50*, 542-546.

Buss, D. M., & Barnes, M. (1986). Preferences in human mate selection. *Journal of Personality & Social Psychology, 50*, 559-570.

Buss, D. M., & Schmitt, D. P. (1993). Sexual strategies theory: An evolutionary perspective on human mating. *Psychological Review, 100*, 204-232.

Buston, P. M., & Emlen, S. T. (2003). Cognitive processes underlying human mate choice: The relationship between self-perception and mate preference in western society. *Proceedings of the National Academy of Sciences, 100*, 8805-8810.

Byrne, D. (1961). Interpersonal attraction and attitude similarity. *The Journal of Abnormal and Social Psychology, 62*, 713-715.

Campbell, L., Chin, K., & Stanton, S. C. (2016). Initial evidence that individuals form new relationships with partners that more closely match their ideal preferences. *Collabra, 2*, 2.

Campbell, L., Simpson, J. A., Boldry, J., & Kashy, D. A. (2005). Perceptions of conflict and support in romantic relationships: The role of attachment anxiety. *Journal of Personality & Social Psychology, 88*, 510-531.

Campbell, L., & Stanton, S. C. (2014). The predictive validity of ideal partner preferences in relationship formation: What we know, what we don't know, and why it matters. *Social and Personality Psychology Compass, 8*, 485-494.

Campbell, L., & Wilbur, C. J. (2009). Are the traits we prefer in potential mates the traits they value in themselves? An analysis of sex differences in the self-concept. *Self and Identity, 8*, 418-446.

Chiao, J. Y., Adams, R. B., Peter, U. T., Lowenthal, W. T., Richeson, J. A., & Ambady, N. (2008). Knowing who's boss: Fmri and erp investigations of social dominance perception. *Group Processes & Intergroup Relations, 11*, 201-214.

Conroy-Beam, D., & Buss, D. M. (2016). Do mate preferences influence actual mating decisions? Evidence from computer simulations and three studies of mated couples. *Journal of Personality and Social Psychology, 111*, 53-66.

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.

Donnellan, M. B., Conger, R. D., & Bryant, C. M. (2004). The big five and enduring marriages. *Journal of Research in Personality, 38*, 481-504.

Durante, K. M., Eastwick, P. W., Finkel, E. J., Gangestad, S. W., & Simpson, J. A. (2016). Pair-bonded relationships and romantic alternatives: Toward an integration of evolutionary and relationship science perspectives. In J. M. Olson & M. P. Zanna (Eds.), *Advances in Experimental Social Psychology (pp.1-74).* Burlington: Academic Press.

Eastwick, P. W. (2016). The emerging integration of close relationships research and evolutionary psychology. *Current Directions in Psychological Science, 25*, 183-190.

Eastwick, P. W., & Buck, A. A. (2014). Too much matching: A social relations model enhancement of the pairing game. *Teaching of Psychology, 241*, 246-250.

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. doi: 10.1037/a0024062

Eastwick, P. W., Keneski, E., Morgan, T. A., & MacDonald, M. A. (2016). *What do short-term and long-term relationships look like? Building the Relationship Coordination and Strategic Timing (ReCAST) model.* Working paper, University of California, Davis, CA. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2820704

Eastwick, P. W., & Hunt, L. L. (2014). Relational mate value: Consensus and uniqueness in romantic evaluations. *Journal of Personality and Social Psychology, 106*, 728-751.

Eastwick, P. W., Luchies, L. B., Finkel, E. J., & Hunt, L. L. (2014a). The predictive validity of ideal partner preferences: A review and meta-analysis. *Psychological Bulletin, 140*, 623-665.

Eastwick, P. W., Luchies, L. B., Finkel, E. J., & Hunt, L. L. (2014b). The many voices of darwin's descendants: Reply to Schmitt (2014). *Psychological Bulletin, 140*, 673-681.

Eastwick, P. W., & Neff, L. A. (2012). Do ideal partner preferences predict divorce? A tale of two metrics. *Social Psychological and Personality Science, 3*, 667-674.

Elder, G. H. (1969). Appearance and education in marriage mobility. *American Sociological Review, 34,* 519-533.

Elliott, D. S., & Morse, B. J. (1989). Delinquency and drug use as risk factors in teenage sexual activity. *Youth and Society, 21,* 32-60.

Ellis, B. J., & Kelley, H. H. (1999). The pairing game: A classroom demonstration of the matching phenomenon. *Teaching of Psychology, 26,* 118-121.

Feingold, A. (1988). Matching for attractiveness in romantic partners and same-sex friends: A meta-analysis and theoretical critique. *Psychological Bulletin, 104,* 226-235.

Festinger, L., Schachter, S., & Back, K. W. (1950). *Social pressures in informal groups: A study of human factors in housing.* New York: Harper & Brothers.

Fink, B., & Penton-Voak, I. (2002). Evolutionary psychology of facial attractiveness. *Current Directions in Psychological Science, 11,* 154-158.

Finkel, E. J., Eastwick, P. W., Karney, B. R., Reis, H. T., & Sprecher, S. (2012). Online dating: A critical analysis from the perspective of psychological science. *Psychological Science in the Public Interest, 13,* 3-66.

Finkel, E. J., Simpson, J. A., Eastwick, P. W. (in press). The psychology of close relationships: Fourteen core principles. *Annual Review of Psychology.*

Fisher, H. E. (1989). Evolution of human serial pair bonding. *American Journal of Physical Anthropology, 73,* 331-354.

Fletcher, G. J. O., & Kerr, P. S. G. (2010). Through the eyes of love: Reality and illusion in intimate relationships. *Psychological Bulletin, 136,* 627-658.

Fletcher, G. J. O., 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., & Thomas, G. (2000a). 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. (2000b). The measurement of perceived relationship quality components: A confirmatory factor analytic approach. *Personality and Social Psychology Bulletin, 26*, 340-354.

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

Gangestad, S. W., & Simpson, J. A. (2000). The evolution of human mating: Trade-offs and strategic pluralism. *Behavioral and Brain Sciences, 23*, 573-644.

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences, 102*, 74-78.

Goodwin, R., Marshall, T., Fülöp, M., Adonu, J., Spiewak, S., Neto, F., & Plaza, S. H. (2012). Mate value and self-esteem: Evidence from eight cultural groups. *PLoS One, 7*, e36106.

Harden, K. P., & Mendle, J. (2011). Adolescent sexual activity and the development of delinquent behavior: The role of relationship context. *Journal of youth and adolescence, 40*, 825-838.

Harris, K. M., & Udry, J. R. (1994-2008). National Longitudinal Study of Adolescent Health (Add Health): Core Files [Restricted Use]. ICPSR27021-v11. Chapel Hill, NC: Carolina Population Center, University of North Carolina-Chapel Hill/Ann Arbor, MI: Inter-

university Consortium for Political and Social Research [distributors], 2013-01-17. http://doi.org/10.3886/ICPSR27021.v11

Hazan, C., & Diamond, L. M. (2000). The place of attachment in human mating. *Review of General Psychology, 4*, 186-204.

Hazan, C., & Shaver, P. R. (1994). Attachment as an organizational framework for research on close relationships. *Psychological Inquiry, 5*, 1-22.

Hunt, L. L., Eastwick, P. W., & Finkel, E. J. (2015). Leveling the playing field: Longer acquaintance predicts reduced assortative mating on attractiveness. *Psychological Science, 26*, 1046-1053.

Huston, T. L., Surra, C. A., Fitzgerald, N. M., & Cate, R. M. (1981). From courtship to marriage: Mate selection as an interpersonal process. In S. Duck & R. Gilmour (Eds.), *Personal relationships (vol. 2): Developing personal relationships* (pp. 53-88). London: Academic Press.

Joel, S., Eastwick, P. W., & Finkel, E. J. (2016). Is romantic desire predictable? Machine learning applied to initial romantic attraction. *Unpublished Manuscript, University of Texas at Austin, Austin, TX*.

Joyner, K., & Udry, J. R. (2000). You don't bring me anything but down: Adolescent romance and depression. *Journal of Health and Social Behavior*, 369-391.

Kaestle, C. E., & Halpern, C. T. (2005). Sexual activity among adolescents in romantic relationships with friends, acquaintances, or strangers. *Archives of pediatrics and adolescent medicine, 159*, 849-853.

Kalick, S. M., & Hamilton, T. E. (1986). The matching hypothesis reexamined. *Journal of Personality and Social Psychology, 51*, 673-682.

Kalmijn, M. (1998). Intermarriage and homogamy: Causes, patterns, trends. *Annual Review of Sociology, 24*, 395-421.

Kalmijn, M., & Flap, H. (2001). Assortative meeting and mating: Unintended consequences of organized settings for partner choices. *Social Forces, 79*, 1289-1312.

Karney, B. R., & Bradbury, T. N. (1995). The longitudinal course of marital quality and stability: A review of theory, methods, and research. *Psychological Bulletin, 118*, 3-34.

Kavanagh, P. S., Robins, S. C., & Ellis, B. J. (2010). The mating sociometer: A regulatory mechanism for mating aspirations. *Journal of Personality and Social Psychology, 99*, 120-132.

Kelley, H. H., Holmes, J. G., Kerr, N. L., Reis, H. T., Rusbult, C. E., & Van Lange, P. A. M. (2003). *An atlas of interpersonal situations*: Cambridge University Press.

Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis*. New York: Guilford.

Kenny, D. A., Albright, L., Malloy, T. E., & Kashy, D. A. (1994). Consensus in interpersonal perception: Acquaintance and the big five. *Psychological Bulletin, 116*, 245-258.

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

Kenrick, D. T., Groth, G. E., Trost, M. R., & Sadalla, E. K. (1993). Integrating evolutionary and social exchange perspectives on relationships: Effects of gender, self-appraisal, and involvement level on mate selection criteria. *Journal of Personality and Social Psychology, 64*, 951-969.

Kenrick, D. T., Neuberg, S. L., Zierk, K. L., & Krones, J. M. (1994). Evolution and social cognition: Contrast effects as a function of sex, dominance, and physical attractiveness. *Personality and Social Psychology Bulletin, 20*, 210-217.

Kirsner, B. R., Figueredo, A. J., & Jacobs, W. J. (2003). Self, friends, and lovers: Structural relations among beck depression inventory scores and perceived mate values. *Journal of affective disorders, 75*, 131-148.

Knapp, M. L., Vangelisti, A. L., & Caughlin, J. P. (2014). *Interpersonal communication & human relationships* (7th ed.): Pearson Higher Ed.

Lakey, B., McCabe, K. M., Fisicaro, S. A., & Drew, J. B. (1996). Environmental and personal determinants of support perceptions: Three generalizability studies. *Journal of Personality and Social Psychology, 70*, 1270-1280.

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.

Langlois, J. H., & Roggman, L. A. (1990). Attractive faces are only average. *Psychological Science, 1*, 115-121.

Lauderdale, B. E., & Linzer, D. (2015). Under-performing, over-performing, or just performing? The limitations of fundamentals-based presidential election forecasting. *International Journal of Forecasting, 31*, 965-979.

Lawrance, K.-a., & Byers, E. S. (1995). Sexual satisfaction in long-term heterosexual relationships: The interpersonal exchange model of sexual satisfaction. *Personal Relationships, 2*, 267-285.

Le, B., Dove, N. L., Agnew, C. R., Korn, M. S., & Mutso, A. A. (2010). Predicting nonmarital romantic relationship dissolution: A meta-analytic synthesis. *Personal Relationships, 17,* 377-390.

Little, A. C., Jones, B. C., & DeBruine, L. M. (2011). Facial attractiveness: Evolutionary based research. *Philosophical Transactions of the Royal Society B: Biological Sciences, 366,* 1638-1659.

Luo, S., & Klohnen, E. C. (2005). Assortative mating and marital quality in newlyweds: A couple-centered approach. *Journal of Personality and Social Psychology, 88,* 304-326.

Luo, S., & Zhang, G. (2009). What leads to romantic attraction: Similarity, reciprocity, security, or beauty? Evidence from a speed dating study. *Journal of Personality, 77,* 933-964.

Lykken, D. T., & Tellegen, A. (1993). Is human mating adventitious or the result of lawful choice? A twin study of mate selection. *Journal of Personality & Social Psychology, 65,* 56-68.

Mahoney, A., Pargament, K. I., Jewell, T., Swank, A. B., Scott, E., Emery, E., & Rye, M. (1999). Marriage and the spiritual realm: The role of proximal and distal religious constructs in marital functioning. *Journal of family psychology, 13,* 321-338.

Maner, J. K., DeWall, C. N., & Gailliot, M. T. (2008). Selective attention to signs of success: Social dominance and early stage interpersonal perception. *Personality and Social Psychology Bulletin, 34,* 488-501.

Maner, J. K., Gailliot, M. T., & DeWall, C. N. (2007). Adaptive attentional attunement: Evidence for mating-related perceptual bias. *Evolution and Human Behavior, 28,* 28-36.

Mare, R. D. (1991). Five decades of educational assortative mating. *American Sociological Review, 56,* 15-32.

Marlowe, F. W. (2004). Mate preferences among Hadza hunter-gatherers. *Human nature, 15,* 365-376.

Martin, N. G., Eaves, L. J., Heath, A. C., Jardine, R., Feingold, L. M., & Eysenck, H. J. (1986). Transmission of social attitudes. *Proceedings of the National Academy of Sciences, 83*, 4364-4368.

Mascie-Taylor, C., & Vandenberg, S. (1988). Assortative mating for iq and personality due to propinquity and personal preference. *Behavior Genetics, 18*, 339-345.

McNulty, J. K. (2013). Personality and relationships. In J. A. Simpson & L. Campbell (Eds.), *The oxford handbook of close relationships*. New York: Oxford University Press.

McNulty, J. K., & Widman, L. (2013). The implications of sexual narcissism for sexual and marital satisfaction. *Archives of sexual behavior, 42*, 1021-1032.

Montoya, R. M., Horton, R. S., & Kirchner, J. (2008). Is actual similarity necessary for attraction? A meta-analysis of actual and perceived similarity. *Journal of Social and Personal Relationships, 25*, 889-922.

Motyl, M., Iyer, R., Oishi, S., Trawalter, S., & Nosek, B. A. (2014). How ideological migration geographically segregates groups. *Journal of Experimental Social Psychology, 51*, 1-14.

Murray, S. L., Holmes, J. G., & Griffin, D. W. (1996). The self-fulfilling nature of positive illusions in romantic relationships: Love is not blind, but prescient. *Journal of Personality and Social Psychology, 71*, 1155-1180.

Murray, S. L., Holmes, J. G., & Griffin, D. W. (2000). Self-esteem and the quest for felt security: How perceived regard regulates attachment processes. *Journal of Personality & Social Psychology, 78*, 478-498.

Newcomb, T. M. (1961). *The acquaintance process*. New York: Holt, Rinehart, & Winston.

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences, 105*, 11087-11092.

Orr, L., Feins, J. D., Jacob, R., Beecroft, E., Sanbonmatsu, L., Katz, L. F., Liebman, J. B. & Kling, J. R. (2003). *Moving to Opportunity: Interim Impacts Evaluation.* Washington D.C.: U.S. Department of Housing and Urban Development.

Overall, N. C., Fletcher, G. J. O., & Simpson, J. A. (2006). Regulation processes in intimate relationships: The role of ideal standards. *Journal of Personality and Social Psychology, 91*, 662-685.

Penke, L., Todd, P. M., Lenton, A. P., & Fasolo, B. (2007). How self-assessments can guide human mating decisions. In G. Geher & G. F. Miller (Eds.), *Mating intelligence: New insights into intimate relationships, human sexuality, and the mind's reproductive system* (pp. 37-75). Mahwah, NJ: Lawrence Erlbaum.

Perrett, D. I., Lee, K., Penton-Voak, I., Rowland, D., Yoshikawa, S., Burt, D., . . . Akamatsu, S. (1998). Effects of sexual dimorphism on facial attractiveness. *Nature, 394*, 884-887.

Perrett, D. I., Penton-Voak, I. S., Little, A. C., Tiddeman, B. P., Burt, D. M., Schmidt, N., . . . Barrett, L. (2002). Facial attractiveness judgements reflect learning of parental age characteristics. *Proceedings of the Royal Society of London B: Biological Sciences, 269*, 873-880.

Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annual Review of Psycholology, 57*, 199-226.

Rizzo, C. J., Daley, S. E., & Gunderson, B. H. (2006). Interpersonal sensitivity, romantic stress, and the prediction of depression: A study of inner-city, minority adolescent girls. *Journal of Youth and Adolescence, 35*, 444-453.

Robins, R. W., Caspi, A., & Moffitt, T. E. (2002). It's not just who you're with, it's who you are: Personality and relationship experiences across multiple relationships. *Journal of Personality, 70*, 925-964.

Rule, N. O., Ambady, N., Adams, R. B., & Macrae, C. N. (2008). Accuracy and awareness in the perception and categorization of male sexual orientation. *Journal of personality and social psychology, 95*, 1019-1028.

Rusbult, C. E., Olsen, N., Davis, J. L., & Hannon, P. A. (2001). Commitment and relationship maintenance mechanisms. In J. H. Harvey & A. Wenzel (Eds.), *Close romantic relationships: Maintenance and enhancement* (pp. 87–113). Mahwah, NJ: Erlbaum.

Scheib, J. E., Gangestad, S. W., & Thornhill, R. (1999). Facial attractiveness, symmetry and cues of good genes. *Proceedings of the Royal Society of London B: Biological Sciences, 266*, 1913-1917.

Schmitt, D. P. (2014). On the proper functions of human mate preference adaptations: Comment on Eastwick, Luchies, Finkel, and Hunt (2014). *Psychological Bulletin, 140*, 666-672.

Schwartz, C. R., & Mare, R. D. (2012). The proximate determinants of educational homogamy: The effects of first marriage, marital dissolution, remarriage, and educational upgrading. *Demography, 49*, 629-650.

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

Smith, C. V. (2007). In pursuit of 'good'sex: Self-determination and the sexual experience. *Journal of Social and Personal Relationships, 24*, 69-85.

Smith, E. R., & Conrey, F. R. (2007). Agent-based modeling: A new approach for theory building in social psychology. *Personality and social psychology review, 11*, 87-104.

Sprecher, S., & Duck, S. (1994). Sweet talk: The importance of perceived communication for romantic and friendship attraction experienced during a get-acquainted date. *Personality and Social Psychology Bulletin, 20,* 391-400.

Sugiyama, L. S. (2005). Physical attractiveness: An adaptationist perspective. In D. M. Buss (Ed.), *The handbook of evolutionary psychology* (pp. 292–343). Hoboken, NJ: John Wiley & Sons, Inc.

Tidwell, N. D., Eastwick, P. W., & Finkel, E. J. (2013). Perceived, not actual, similarity predicts initial attraction in a live romantic context: Evidence from the speed-dating paradigm. *Personal Relationships, 20,* 199-215.

Townsend, J. M. (1989). Mate selection criteria: A pilot study. *Ethology & Sociobiology, 10,* 241-253.

Walster, E., Aronson, V., Abrahams, D., & Rottmann, L. (1966). Importance of physical attractiveness in dating behavior. *Journal of Personality and Social Psychology, 4,* 508-516.

Watson, D., Klohnen, E. C., Casillas, A., Nus Simms, E., Haig, J., & Berry, D. S. (2004). Match makers and deal breakers: Analyses of assortative mating in newlywed couples. *Journal of Personality, 72,* 1029-1068.

Weigel, D., & Murray, C. (2000). The paradox of stability and change in relationships: What does chaos theory offer for the study of romantic relationships? *Journal of Social and Personal Relationships, 17,* 425-449.

Zietsch, B. P., Lee, A. J., Sherlock, J. M., & Jern, P. (2015). Variation in women's preferences regarding male facial masculinity is better explained by genetic differences than by previously identified context-dependent effects. *Psychological Science, 26,* 1440-1448.

Zietsch, B. P., Verweij, K. J., Heath, A. C., & Martin, N. G. (2011). Variation in human mate choice: Simultaneously investigating heritability, parental influence, sexual imprinting, and assortative mating. *The American Naturalist, 177*, 605-616.

Footnotes

[1] The assessment strategy used in the present article examines clustering one trait at a time, so a low level of clustering would mean that factors that are consistent across time have weak effects *on that particular attribute*. Moreover, if variables that are largely or completely unassociated with mating (e.g., day of the week someone was born, middle initial) failed to exhibit clustering, this finding would have negligible implications for the theoretical perspectives discussed here. For this reason, all variables examined in this article have been linked extensively to human mating in one or more literatures in prior research (Supplementary Table 1).

[2] We wish to thank Daniel Conroy-Beam who provided us with the R-code that we then used to create these simulations.

[3] In our view, although mate preferences and similarity-attraction effects could in principle account for the current clustering findings as well, these explanations remain less plausible given that the effect sizes detected here vastly exceed the sizes typically found in direct tests of these two phenomena in large studies and meta-analyses (e.g., Eastwick et al., 2014a; Luo & Zhang, 2009).

[4] We initially identified twenty-two measures, fifteen of which were measured by single-item constructs. Reviewers noted that low intraclass correlations for these variables could be due to the indeterminate reliability of the items. Thus, the current manuscript consists only of the seven multi-item constructs; the method and results section for the previous version of the manuscript can be found here.

[5] As in Study 1, we tested whether the findings differed for men and women, but none of the 14 values listed in Table 2 (i.e., focal person percentages with and without controlling for school) significantly differed by participant sex.

[6] When the woman had completed the evaluation form, the website (inexplicably) transformed all 1-star ratings into a 4, 2-star ratings into a 6.5, 3-star ratings into an 8, 4-star ratings into a 9, and 5-star ratings into a 10. We transformed the numbers obtained from the website back to the original 1-5 metric for analyses.

[7] Unfortunately, this dataset does not permit us to easily replicate the analyses from Studies 1 and 2 that subtracted casual dating partners (although these analyses revealed identical conclusions in Studies 1 and 2). Dropping the hookup partners reduced the usable dataset to only $N = 24$ focal men (i.e., only 24 men had multiple women report on them who identified as together or ex-boyfriend). Nevertheless, analyses on these reports alone revealed the same one-factor romantic desirability construct, and importantly, none of the ICCs were significantly different from zero.

[8] Because romantic desirability contained more than two items, we actually ran two versions of this construct using two different randomly generated pairs of bins (as in Eastwick & Hunt, 2014). The two versions differed a mere 1.6%, so we averaged across the two versions in the analyses reported here.

[9] If we instead add the friends to the main analyses, analyses on this larger sample ($N = 367$ focal men; $N = 1445$ female partners/friends) revealed clustering of 7.2% for romantic desirability, 11.2% for positive qualities, and 5.8% for negative qualities.

Table 1

Percentage of Variance in Romantic Partner Qualities (Study 1) Accounted for by Focal Person

Quality	Focal person variance	Similarity correlation
Attractiveness	26.8%	.30***
Masculinity	30.7%	-.16†
Dominance	29.9%	.28***

Note: Analyses control for sex.

Table 2

Percentage of Variance in Romantic Partner Qualities (Study 2) Accounted for by Focal Person (All Waves)

Quality	Focal person variance (no controls)	Focal person variance (school control)	School variance	Similarity correlation	Similarity correlation (school control)
Vitality	6.5%	4.4%	2.7%	.08**	.07*
Depression	9.6%	5.5%	5.1%	.08**	.06*
Delinquency	11.1%	8.1%	3.0%	.05*	.04
Religiosity	11.3%	0.0%	13.8%	.19***	.09**
Educational aspirations	11.5%	4.6%	11.3%	.16***	.12***
Self-esteem	12.4%	9.8%	3.3%	.06*	.04
Intelligence (IQ)	27.5%	4.0%	24.8%	.32***	.18***

Note: Analyses control for sex. Similarity correlation analysis does not control for school.

Table 3.

Percentage of Variance in Romantic Partner Qualities Accounted for by Focal Person (Wave III eliminated)

Quality	Focal person variance (no controls)	Focal person variance (school control)	School variance
Depression	5.5%	2.2%	5.2%
Vitality[a]	6.5%	4.4%	2.7%
Self-esteem	8.7%	6.1%	3.6%
Educational aspirations[a]	11.5%	4.6%	11.3%
Delinquency	16.8%	13.2%	3.5%
Religiosity	26.8%	5.4%	25.6%
Intelligence (IQ)[a]	27.5%	4.0%	24.8%

[a]Row is identical to Table 2 because the measure was not assessed at Wave III

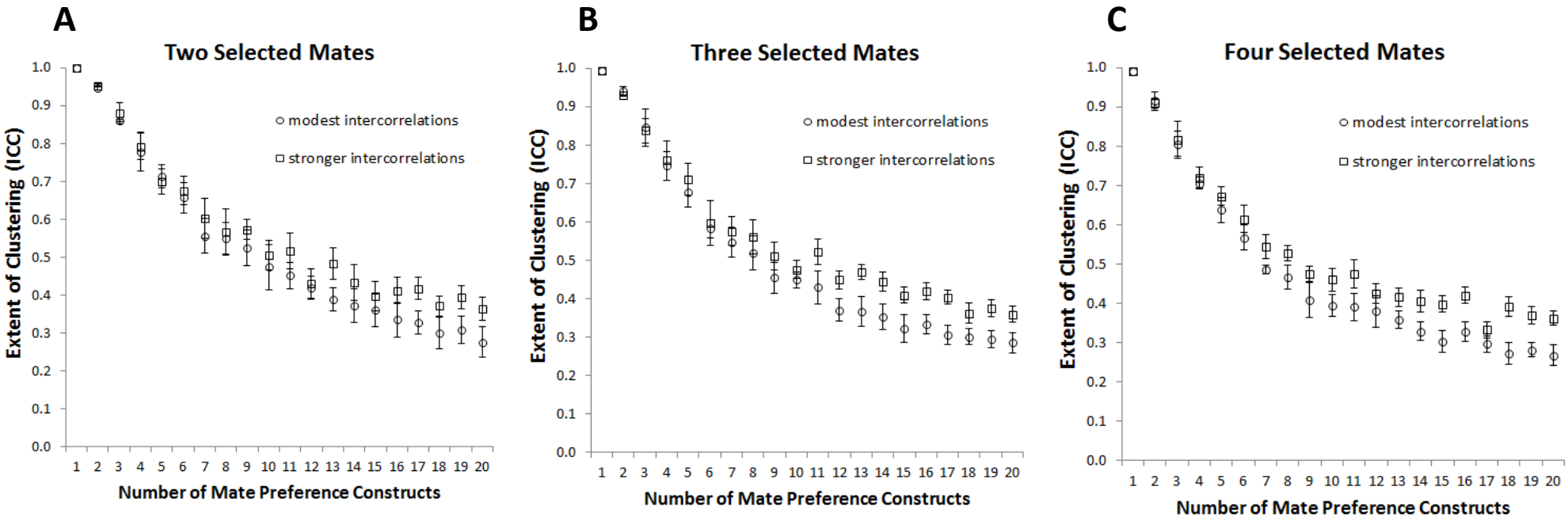
Note: Analyses control for sex.

Table 4

Percentage of Variance in Women's Ratings of Men (Study 3) Accounted for by the Man

Quality	Focal person variance (no controls)	Focal person variance (school control)	School variance
Romantic desirability	8.6%	8.6%	0.0%
Sexual satisfaction	5.7%	5.7%	0.0%
Number of best qualities	7.3%	7.3%	0.0%
Number of worst qualities	3.2%	3.2%	0.0%

Figure 1 – Extent of Clustering in Agent-based Simulations



Note: Agent-based simulations depict the amount of clustering on a given attribute when agents select mates based on the extent to which mates match their preferences (i.e., ICC for the trait across the selected mates). Simulations depict ICCs when two (Panel A), three (Panel B), and four (Panel C) mates are selected. Modest intercorrelations: preferences for traits correlate r = .20; perceptions of mates' traits correlate r = .20. Stronger intercorrelations: preferences for traits correlate r = .20; perceptions of mates' traits correlate r = .45; see text for real-world justification of these values. Error bars depict 95% confidence interval around the predicted ICC.

Supplementary Material: Justification for Preference and Trait Perception Intercorrelation Values in the Pilot Study Simulations

In order to make sure that we built the data structure in our simulations on a real-world empirical foundation, we turned to the largest published dataset we possessed that contained measures of ideal partner preferences and perceptions of a partner's traits. In this study (Eastwick et al., 2011, Study 3), 502 participants (M_{age} = 40.9 years old) provided ratings of ideal partner preferences for 48 traits and rated a current or most desired romantic partner on those same 48 traits. In this dataset, ideal partner preference ratings correlated with other ideal partner preference ratings at approximately r = .20 on average, and partner trait ratings correlated with other partner trait ratings at approximately r = .20 on average. These correlations were nearly identical whether we calculated (a) the absolute value of the correlation among the 7 factors that comprise the 48 individual items or (b) the absolute value of the correlation among all 48 items themselves. This r = .20 value is likely conservative; some reports find considerably higher correlations among ideal partner preference ratings (e.g., ranging from r = .24 to r = .66; Lam et al., 2016). Thus, we incorporated into our model this r = .20 correlation among ideal preference judgments and among partner trait judgments ("modest intercorrelations" simulations). We also report a set of simulations ("stronger intercorrelations") where we raise the intercorrelation among the trait perceptions (but not the ideal preference judgments) to r = .45 to reflect initial impression contexts: Perceptions of traits correlate with other perceptions of traits at r = .45 in the speed-dating research of Eastwick and Finkel (2008).

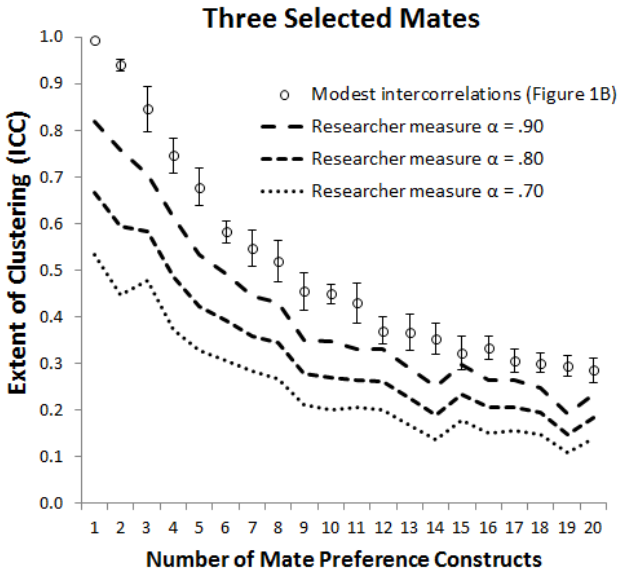
Supplementary Material: Incorporating Reliability of Trait Measurement into Pilot Study Simulations

The ICC for a given attribute is likely to be affected by the reliability of the measure that researchers use to assess it (i.e., in a population of selected mates). The figures below are adaptations of the pilot study simulations with three mates (i.e., Figure 1B) that address the question: How would reliability affect the ICC that researchers would be able to detect in a world where people select mates based on their preferences for particular attributes?

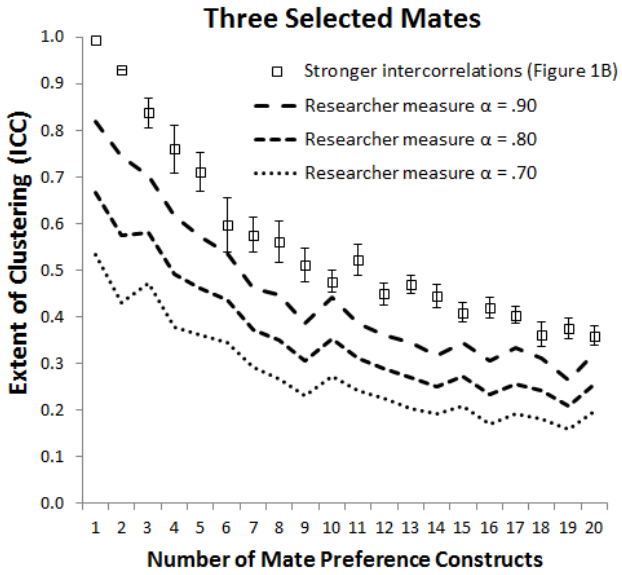
These figures depict the ICCs that emerge when the ICCs are calculated not on the (exact) attributes of the selected mates but rather on "jiggled" versions of these attributes. Specifically, we added random noise to the attributes in three different simulations such that they correlated with the actual attribute scores at r = .90, .80, and .70 (to simulate a measure with α = .90, .80, and .70, respectively). We conducted these simulations for both the "modest intercorrelations" values (i.e., preferences for traits correlate r = .20; perceptions of mates' traits correlate r = .20) and the "stronger intercorrelations" values (i.e., preferences for traits correlate r = .20; perceptions of mates' traits correlate r = .45).

As depicted in Figure S1 (modest intercorrelations) and S2 (stronger intercorrelations), increases in reliability boost the ICC that researchers can detect. Intriguingly, reliability had the most dramatic impact when the agents selected mates based on a small number of attributes; as the number of attributes increased, the effect of reliability was less pronounced. In all cases, the detectable ICC remained greater than the "meaningful" threshold of .10 (Kenny et al., 2006). Nevertheless, researchers will have the greatest success obtaining stable estimates of ICCs to the extent that they are using reliable measures.

S1



S2



Supplementary Material: Add Health Item Codes for Multi-item Constructs

Wave 1

- Delinquency
  - Delinquency = sum(H1DS1, H1DS2, H1DS3, H1DS4, H1DS5, H1DS6, H1DS7, H1DS8, H1DS9, H1DS10, H1DS11, H1DS12, H1DS13, H1DS14, H1DS15);
  - lnDelinquency = log(Delinquency+1);
- Depression
  - H1FS4R = 3 - H1FS4;
  - H1FS8R = 3 - H1FS8;
  - H1FS11R = 3 - H1FS11;
  - H1FS15R = 3 - H1FS15;
  - Depression = mean(H1FS1, H1FS2, H1FS3, H1FS4R, H1FS5, H1FS6, H1FS7, H1FS8R, H1FS9, H1FS10, H1FS11R, H1FS12, H1FS13, H1FS14, H1FS15R, H1FS16, H1FS17, H1FS18, H1FS19);
- Educational Aspirations = mean(H1EE1, H1EE2);
- Vitality = mean(H1PF26, H1PF31);
- Religiosity
  - H1RE3New = H1RE3;
  - if H1RE1 = 0 then H1RE3New = 4;
  - if H1RE3 = 9 then H1RE3New = .;
  - H1RE4New = H1RE4;
  - if H1RE1 = 0 then H1RE4New = 4;
  - H1RE6New = H1RE6;
  - if H1RE1 = 0 then H1RE6New = 5;
  - Religiosity = mean(H1RE3New, H1RE4New, H1RE6New);
- Self-esteem = mean(H1PF32, H1PF36, H1PF33, H1PF30, H1PF35, H1PF34);
- Parental Education
  - PA12New = PA12;
  - if PA12 = 4 then PA12New = 3;
  - if PA12 = 5 then PA12New = 3;
  - if PA12 = 6 then PA12New = 4;
  - if PA12 = 7 then PA12New = 5;
  - if PA12 = 8 then PA12New = 6;
  - if PA12 = 9 then PA12New = 7;
  - if PA12 = 10 then PA12New = .;
  - PB8New = PB8;
  - if PB8 = 4 then PB8New = 3;
  - if PB8 = 5 then PB8New = 3;
  - if PB8 = 6 then PB8New = 4;
  - if PB8 = 7 then PB8New = 5;
  - if PB8 = 8 then PB8New = 6;
  - if PB8 = 9 then PB8New = 7;
  - if PB8 ge 10 then PB8New = .;

    - proc standard mean = 0 SD = 1; var PB8New PA12New; run;
    - ParentalEducation = mean(PB8New, PA12New);

Wave 2

- Delinquency
    - Delinquency = sum(H2DS1, H2DS2, H2DS3, H2DS4, H2DS5, H2DS6, H2DS7, H2DS8, H2DS9, H2DS10, H2DS11, H2DS12, H2DS13, H2DS14);
    - lnDelinquency = log(Delinquency+1);
- Depression
    - H2FS4R = 3 - H2FS4;
    - H2FS8R = 3 - H2FS8;
    - H2FS11R = 3 - H2FS11;
    - H2FS15R = 3 - H2FS15;
    - Depression = mean(H2FS1, H2FS2, H2FS3, H2FS4R, H2FS5, H2FS6, H2FS7, H2FS8R, H2FS9, H2FS10, H2FS11R, H2FS12, H2FS13, H2FS14, H2FS15R, H2FS16, H2FS17, H2FS18, H2FS19);
- Educational Aspirations = mean(H2EE1, H2EE2);
- Vitality = mean(H2PF17, H2PF22);
- Religiosity
    - H2RE3New = H2RE3;
    - if H2RE1 = 29 then H2RE3New = 4;
    - H2RE4New = H2RE4;
    - if H2RE1 = 29 then H2RE4New = 4;
    - H2RE6New = H2RE6;
    - if H2RE1 = 29 then H2RE6New = 5;
    - Religiosity = mean(H2RE3New, H2RE4New, H2RE6New);
- Self-esteem = mean(H2PF23, H2PF27, H2PF24, H2PF21, H2PF26, H2PF25);

Wave 3

- Delinquency
    - Delinquency = mean(H3DS1, H3DS2, H3DS3, H3DS4, H3DS5, H3DS6, H3DS7, H3DS8, H3DS9, H3DS10, H3DS11, H3DS12, H3DS13);
    - lnDelinquency = log(Delinquency+1);
- Depression
    - H3SP7R = 3 -H3SP7;
    - H3SP11R = 3 - H3SP11;
    - Depression = mean(H3SP13, H3SP12, H3SP11R, H3SP10, H3SP9, H3SP8, H3SP7R, H3SP6, H3SP5);
- Self-esteem = mean(H3SP19, H3SP20, H3SP21, H3SP22);
- Religiosity = mean(H3RE24, H3RE30, H3RE32);

Supplementary Table 1 – Exemplar Studies Examining Variables of Interest

Quality	Three Example Studies
Study 1	
Attractiveness	Kenrick et al., 1994; Langlois & Roggman, 1990; Scheib et al., 1999
Masculinity	Burriss et al., 2011b; Perrett et al., 1998; Zietsch et al., 2015
Dominance	Bailey et al., 2011; Chiao et al., 2008; Maner et al., 2008
Study 2	
Vitality	Campbell & Wilbur, 2009; Fletcher et al., 2014; Overall et al., 2006;
Religiosity	Luo & Klohnen, 2005; Mahoney et al., 1999; Martin et al., 1986
Educational aspirations	Elder, 1969; Schwartz & Mare, 2012; Townsend, 1989
Delinquency	Armour & Haynie, 2007; Belsky et al., 2010; Elliot & Morse, 1989
Self-esteem	Bale & Archer, 2013; Goodwin et al., 2012; Murray et al., 2000
Depression	Joyner & Udry, 2000; Kirsner et al., 2003; Rizzo et al., 2006
Intelligence	Buss & Barnes, 1986; Walster et al., 1966; Watson et al., 2004
Study 3	
Romantic desirability	Campbell et al., 2016; Fletcher et al., 2000a; Murray et al., 1996
Sexual satisfaction	Lawrance & Byers, 1995; McNulty & Widman, 2013; Smith, 2007
Number of best/worst qualities	(none)

Note: Cited articles use the same assessment strategies as the current studies. That is, Study 1 exemplar studies used ratings from photographs, Study 2 exemplar studies used ratings from self-reports or cognitive tests (i.e., for *intelligence*), Study 3 exemplar studies used target-specific ratings of positive traits (i.e., for *romantic desirability*) and positive sexual experiences (i.e., for *sexual satisfaction*). All the studies included here examine the relevant construct in the context of romantic relationships and/or evolutionary psychological processes. (We could locate no prior studies which operationalized positivity/negativity as the number of qualities selected from a list of possible qualities.)

Supplementary Table 2 – Study 2 Ratings for Seven Attributes

Quality	General	High school
Delinquency	-5.35	-3.18
Depression	-2.03	-1.86
Religiosity	1.04	0.87
Intelligence (IQ)	4.13	3.13
Educational aspirations	4.54	4.80
Vitality	4.92	5.65
Self-esteem	4.98	4.45

Note: Scale ranged from -9 (*no one thinks this person would be a valuable mate*) to 0 (*some people think this person would be a valuable mate*) to 9 (*everyone thinks this person would be a valuable mate*).

Supplementary Table 3 – Percentage of Variance in Romantic Partner Qualities Accounted for by Focal Person (Sexual Partners Removed)

Quality	Focal person variance (no controls)	Focal person variance (school control)	School variance
Vitality	5.5%	3.1%	2.8%
Religiosity	6.2%	0.0%	9.2%
Educational aspirations	7.3%	0.9%	12.1%
Delinquency	8.5%	6.2%	2.0%
Self-esteem	11.3%	10.0%	1.7%
Depression	11.4%	7.2%	5.2%
Intelligence (IQ)	28.9%	8.3%	22.6%

Note: Analyses control for sex.

Supplementary Table 4 – Percentage of Variance in Romantic Partner Qualities Accounted for by Focal Person (Wave I only)

Quality	Focal person variance (no controls)	Focal person variance (school control)	School variance
Depression	7.4%	5.6%	2.5%
Self-esteem	8.2%	8.2%	0.0%
Vitality	12.4%	11.2%	1.6%
Delinquency	13.3%	11.1%	2.5%
Educational aspirations	13.7%	8.3%	8.4%
Intelligence (IQ)	26.1%	7.6%	20.8%
Religiosity	28.2%	8.1%	23.3%

Note: Analyses control for sex.

Supplemental Table 5 – Percentage of Variance in Romantic Partner Qualities Accounted for by Focal Person (Wave II only)

Quality	Focal person variance (no controls)	Focal person variance (school control)	School variance
Self-esteem	0.0%	0.0%	0.0%
Vitality	15.6%	15.6%	0.0%
Depression	19.4%	19.4%	0.0%
Religiosity	22.1%	6.6%	22.5%
Intelligence (IQ)	29.7%	0.0%	44.9%
Delinquency	29.9%	27.0%	3.3%
Educational aspirations	34.4%	15.3%	25.1%

Note: Analyses control for sex.