Re-evaluating the Influence of Preferences on Attraction and Relationship Decisions

Dissertation

zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades

"Doctor rerum naturalium"

der Georg-August-Universität Göttingen

im Promotionsstudiengang Behavior and Cognition

der Georg-August University School of Science (GAUSS)

vorgelegt von

Julie Christin Driebe

aus Bad Hersfeld

Göttingen, 2021

Danksagung

Ich möchte all den Menschen von ganzem Herzen danken, die mich in den letzten Jahren auf dem Weg zu meiner Dissertation unterstützt haben.

Mein besonderer gilt meinem Doktorvater Lars Penke, der mir ermöglicht hat dieses Projekt zustande zu bringen. Außerdem möchte ich ihm für seine Unterstützung und Anregungen zu neuen Ideen danken. Von seiner Erfahrung und Expertise durfte ich mehr als einmal profitieren. Ebenso möchte ich Tanja Gerlach danken, ohne deren Unterstützung und Betreuung ich diese Dissertation wahrscheinlich nie hätte fertigstellen können.

Meinem Thesis Committee, Mitja Back und Julia Ostner, möchte ich für das wertvolle Feedback und das kritische Hinterfragen danken, das mir geholfen hat neue Blickwinkel einzunehmen.

Tanya Behne, Maragrete Boos und Anne Schacht danke ich für die Bereitschaft meinem Prüfungskomitee anzugehören.

Christoph, Lara, Laura und Tobi möchte ich für die unzähligen schönen und lustigen Momente danken, die meinen Arbeitsalltag und meine Freizeit bereichert haben. Nicht zuletzt wart ihr es, die diese Arbeit durch ihre Mithilfe und durch ihren Zuspruch ermöglicht haben. Ruben danke ich für die unzähligen Male, die er mich seit meiner Bachelorarbeit unterstützt hat. Und Julia danke ich für ihre seelische und tatkräftige Unterstützung, die mich bis zu den letzten Metern meiner Promotion immer wieder motiviert hat.

I would like to thank Paul, Michi, Juliane, Morgan and Brendan for their great collaboration.

Anna, Chiara, Daniel, Dorle, Fenja, Johanna, Paula Bange, Paula Tönshoff, Lara, Lena, Louisa, Marieke und Silvi danke ich für die großartige Unterstützung bei meiner Datenerhebung und in meiner Lehre.

Auch meinen Freunden, insbesondere AC, Anja, Britta, Felix, Hubi, Jana, Laura, Lukas, Maike, Maria und Missi möchte danken, dass sie mir auf unserem gemeinsamen Weg seit Kindertagen zur Seite standen. Sie haben mir immer wieder den nötigen Ausgleich gegeben und mich so in dieser herausfordernden Zeit unterstützt.

Meiner Familie, insbesondere meinen Eltern und meinem Bruder Christopher, danke ich von ganzem Herzen für ihre grenzenlose und liebevolle Unterstützung und dass sie jederzeit für mich da sind. Und zuletzt möchte ich Paul danken, für seine Liebe, Geborgenheit und Aufmunterungen, die mir unglaublich viel Kraft schenken.

Table of Contents

Appendix A. Manuscript 1 (Intelligence can be detected but is not found attractive)

Appendix B. Manuscript 2 (Ideal Partner Preferences are Stable Over 13 Years)

Appendix C. Manuscript 3 (Ideals are Associated with Perceived Traits of Future Partners across 13 Years, With a Closer Match Being Related to Commitment)

Appendix D. Curriculum Vitae

Preface

This dissertation is not a cumulative, publication-based dissertation, but follows it in form. It includes three manuscripts, one of which having been accepted for publication, one being submitted for publication and one being in preparation for submission.

Driebe, J.C.*, Sidari, M.J.*, Dufner, M., von der Heiden, J. M., Bürkner, P.C., Penke, L.**, Zietsch, B.P.**, & Arslan, R.C.** (2021). Intelligence can be detected but is not found attractive. *Evolution and Human Behavior*. Advance online publication. https://doi.org/10.1016/j.evolhumbehav.2021.05.002.

      *Julie Driebe and Morgan Sidari share the first authorship

      **Lars Penke, Brendan Zietsch and Ruben Arslan share the last authorship

Driebe, J.C., Stern, J., Penke, L., Gerlach, T.M. (2021). *Ideal Partner Preferences are Stable Over 13 Years* [Manuscript submitted for publication]. Department of Biological Personality Psychology, University of Goettingen.

Driebe, J.C., Stern, J., Penke, L., Gerlach, T.M. (2021). *Ideals are Associated with Perceived Traits of Future Partners across 13 Years, With a Closer Match Being Related to Commitment* [Manuscript in preparation]. Department of Biological Personality Psychology, University of Goettingen.

In the present thesis, I explain the theoretical background connecting three manuscripts which lead to the question I aim to answer with this thesis. These articles are attached in the Appendix (A, B, C). I serve as the first author in all three manuscripts (shared first authorship in manuscript A). The following table indicates in which part I was involved in each manuscript. Please note, that in each process my co-authors were involved as well, meaning that each step is not my sole contribution but the result of a shared project.

	Manuscript 1	Manuscript 2	Manuscript 3
Study Design		X	X
Data Collection	X	X	X
Data Analyses	X	X	X
Writing the Manuscript	X	X	X

I hereby declare that all parts of this dissertation were written by myself, that assistance of third parties was only accepted if scientifically justifiable and acceptable in regards to the examination regulations, and that all sources have been quoted.

Hannover, 2021   *J.D.*

0.   Abstracts

0.1   English Abstract

Several theories assume that humans possess partner preferences influencing their selection of a partner. It has been proposed that humans seek a partner meeting their preferences and that a match between their preferences and a partner's characteristics influences a relationship's well-being. However, the evidence supporting this hypothesis is mixed, which challenges the idea that preferences guide decisions regarding the relationship (starting from preferences influencing decisions of initial attraction up to established relationships).

In this dissertation, I therefore describe four reasons which could explain the ambiguous findings of previous research (unrepresentative samples, study design as a confounding factor, analytical approaches, and investigated time frame). With my three complementing manuscripts, I aimed to address each of these shortcomings and re-evaluate the influence of preferences on relationship decisions. In Manuscript 1, my co-authors and I found that more intelligent individuals are not rated as being more appealing; mainly physical attractiveness seemed to be decisive for initial attraction. In Manuscript 2, we found that partner preferences are considerably stable over a period of 13 years. At the same time, individuals possessed only to a certain extend insight into how their own partner preferences have changed over this period of time. In Manuscript 3, we found that partner preferences are related to perceived characteristics of partners in the subsequent 13 years. A closer match between preferences and perceived partner characteristics was associated with higher relationship commitment. While I discussed whether all manuscripts adequately addressed previous limitations, I concluded that preferences influence relationship decisions. I ended with integrating our findings into a model which is able to explain the reduced influence of preferences at initial attraction with an increasing influence at later stages of relationship formation. While I

cannot preclude that other models could also explain the pattern of results, I describe a possible design for future studies, which could investigate how preferences are integrated into relationship decisions.

0.2 Deutsche Zusammenfassung

Eine Vielzahl von Theorien nimmt an, dass Menschen Partnerpräferenzen besitzen, die ihre Partnerwahl beeinflussen. Existierende Theorien postulieren, dass sich Menschen passend zu ihren Präferenzen einen Partner suchen und dass eine Passung zwischen Präferenzen und den Eigenschaften eines Partners die Zufriedenheit der Beziehung beeinflussen. Jedoch sind die Befunde, die diese Hypothese stützen, gemischt. Daher ist es fraglich, ob Präferenzen tatsächlich Entscheidungen beeinflussen, die auf eine Beziehung einwirken (beginnend mit dem Einfluss von Präferenzen auf die anfängliche Anziehung bis hin zu bestehenden Beziehungen). In dieser Dissertation arbeite ich daher vier Gründe heraus, die die uneindeutigen Befunde bisheriger Literatur erklären könnten (nicht repräsentative Stichproben, Studiendesign als Konfundierung, abweichende analytische Ansätze, untersuchte Zeitspanne). Mit meinen drei sich komplementierenden Manuskripten habe ich das Ziel, diese Defizite zu adressieren und zu re-evaluieren, ob Präferenzen Beziehungsentscheidungen beeinflussen. In Manuskript 1, fanden meine Ko-Autoren und ich heraus, dass intelligentere Menschen nicht als anziehender bewertet werden; Hauptsächlich die physische Attraktivität schien für die initiale Anziehung entscheidend zu sein. In Manuskript 2 haben wir herausgefunden, dass Partnerpräferenzen über eine Zeitspanne von 13 Jahren beachtlich stabil sind. Gleichzeitig haben Individuen aber nur teilweise eine Einsicht darüber, wie sich ihre eigenen Partnerpräferenzen über diese Zeit hinweg verändert haben. In Manuskript 3 haben wir herausgefunden, dass Partnerpräferenzen mit den wahrgenommenen Eigenschaften von Partnern in den folgenden 13 Jahren zusammenhängen.

Eine größere Passung zwischen Präferenzen und wahrgenommenen Partnereigenschaften hing mit einer höheren Verbindlichkeit der Beziehung zusammen. In der Diskussion, ob alle Manuskripte vorherige Einschränkungen adäquat adressiert haben, komme ich zu dem Schluss, dass Präferenzen Beziehungsentscheidungen beeinflussen. Ich ende mit der Einordnung der Befunde in ein Modell, welches den reduzierten Einfluss von Präferenzen für die erste Anziehung mit einem steigenden Einfluss von Präferenzen auf spätere Stadien der Entstehung einer Beziehung, erklären kann. Da ich nicht ausschließen kann, dass dieses Ergebnismuster auch durch andere Modelle erklärt werden kann, zeige ich Wege für zukünftige Studien auf, die untersuchen können, wie Präferenzen in Beziehungsentscheidungen integriert werden.

1. Introduction

Who we seek and find as a romantic partner has crucial consequences for our health and well-being (Pietromonaco & Collins, 2017). For example, we know that the life-expectancy of married couples is higher compared to unmarried individuals (Jia & Lubetkin, 2020). As humans spend many years of their lifetime in these romantic relationships, it does not come to a surprise that people strive to find their ideal partner and spend considerable time and effort into this search for a partner. This importance of finding a partner is partly reflected in the flourishing market of dating platforms that promise to help create a happy and well-functioning relationship by finding a perfectly matching partner[1]. On many of these platforms, users create profiles in which they state their interests, personal information and what they are looking for in a partner. Based on this information, the platforms' users receive suggestions of potential partners and can actively search for a suitable partner. This process builds on the assumptions that 1) people are aware of their preferences driving their partner selection and that 2) a match between preferences and a partner's traits is related to a satisfying relationship. The evidence for both assumptions, however, is mixed in psychological research (e.g. Eastwick et al., 2014; Eastwick et al., 2019; Fletcher et al., 2020). So are slogans of dating platforms really build on solid science or just a promising but misleading lie?

This brings about the substantial question this dissertation addresses: Are partner preferences, indeed, related to mate choices at different time points of relationship formation? To this end, this dissertation compromises three manuscripts. The first manuscript focuses on the role of preferences at the initial stage of getting to know a potential partner. Because intelligence is among to the most highly valued characteristics in a partner (e.g. Buss & Barnes, 1986; Walter et al., 2020), we investigated the influence of intelligence on attraction

---

[1] This assumption is based on personal review of the most widely used dating websites.

in particular. With the two following manuscripts we focused on the influence of partner preferences at later time points of relationship formation: We investigated the stability of preferences over 13 years and their influence on partner selection as well as the functioning of relationships over this period of time.

1.1. Theories on Human Mate Choice

Across several fields of psychology, theories developed which aim to explain how people select their partners[2] (for reviews see Conroy-Beam, 2021; Finkel & Baumeister, 2010). These theories are not necessarily mutually exclusive and partly even overlap.

In social psychology, for example, proximity and similarity have been discussed to influence partner selection. Closer proximity enables more interaction between individuals which is associated with higher attraction towards another person (Byrne, 1961; Newcomb, 1956). These findings are related to the *mere exposure effect*, as repeated contact increases positive feelings to another object (Zajonc, 1968), in this case to another person. Moreover, individuals also receive rewards of their interactions with others. For example, when others are perceived to be similar, this is thought to be conceived as a validation of one's own beliefs and attitudes resulting in a positive feeling (Byrne & Blaylock, 1963). Based on this, the *similarity-attraction hypothesis* was developed, proposing that individuals are attracted towards those who are similar. Supporting this, research has found that couples resemble each other regarding several characteristics; this effect is also called *positive assortment* (Lykken & Tellegen, 1993). For example, couples are similar concerning attributes like physical appearance, attitudes, intelligence and personality (Luo, 2017; Watson et al., 2004). Nevertheless, similarity cannot necessarily be ascribed to an active selection of a similar

---

[2] In the following, the terms partner selection, mate choice and mate selection are used interchangeably and all refer to the choice for an individual as a partner.

partner (for an overview see Watson et al., 2014). Instead, individuals could also become more similar over time without necessarily resembling each other at the beginning, an effect which is referred to as *convergence* (Gonzaga et al., 2010). Other reasons could be that partners meet in social environments in which individuals already are similar to each other due to social homogamy (Kalmijn, 1998). In line with this approach, Lykken and Tellegen (1993) argue that similarity coefficients that were found between partners are too small in order for them to be decisive in mate selection and that choices rather seem random within one's social environment (but see Gonzaga et al., 2010; Watson et al., 2004).

Yet, from an evolutionary perspective, humans, as a sexually reproducing species, should have evolved strategies guiding their mate choices with the aim of increasing their *reproductive success* (Buss & Schmitt, 2019; Conroy-Beam & Buss, 2016; Miller & Todd, 1998). Reproductive success relates to the increased chances of survival either directly through, for example, the provision of resources or protection, or indirectly through genetic benefits which are passed onto the next generation (Gangestad & Scheyd, 2005). This assumption dates back to Darwin's theory of sexual selection (Darwin, 1871), explaining the evolution of adaptations not only because of a survival benefit but also because of an advantage in mating (Buss & Schmitt, 2019). Since then, several strategies guiding human mate choice have been proposed which differ for example with regard to contexts (*short-term* [i.e. sexual affairs] and *long-term mating* [i.e. committed relationships]), between the sexes, and different environments (for a review see Buss & Schmitt, 2019). The assumed underlying mechanism of these strategies are mate preferences which are presumed to guide an individual in finding a suitable partner (Buss & Schmitt, 1993). The perspective that human mate choice is based upon principles of evolutionary adaptivity is further supported by research finding that, cross-culturally, women, as compared to men, prefer a long-term partner who has a higher ability to acquire resources while men, as compared to women,

prefer a partner for a committed relationship who is younger and physically more attractive (Buss, 1989; Walter et al., 2020). In evolutionary psychology it is assumed that these sex differences in human mate preferences have evolved to maximise reproductive success (Buss & Barnes, 1986). For example, particularly for men, women's age is a relevant cue, since women can only get pregnant during a certain age. Physical attractiveness is further assumed to indicate a woman's age and health, thus, serving as a further cue of her *reproductive value* (i.e. the expected number of her offspring) (Buss, 1989; Symons, 1979). Because men are not restricted to a reproductive age, these cues are less relevant to women when selecting a partner. For women, a partner with high access to resources, yields immediate but also future benefits for her and her offspring (e.g. higher social status). This should be especially relevant to women as compared to men (Buss, 1989): Since women have a higher minimal parental investment (Trivers, 1972), with an investing partner, women could reduce this disparity. Although different social roles of society may have also led to sex differences (W. Wood & Eagly, 2012), recent findings challenged this theory as sex differences where not smaller in cultures with higher gender equality (Walter et al., 2020; Zhang et al., 2019). In summary, most theories across several disciplines assume that the selection of a partner is not random but the result of lawful choices guided by certain preferences.

1.2. The Influence of Preferences

In the 1940s, Christensen (1947) assessed which attributes individuals wished for in a spouse. He found that women and men alike pay attention to similar characteristics such as the desire for children, ambition or intelligence. Since then, many studies followed investigating what characteristics humans prefer in a partner. This work yielded a large correspondence in reported partner preferences which could be replicated cross-culturally (e.g. Buss, 1989; Buss & Barnes, 1986; Walter et al., 2020) and across diverse research methods (Li et al., 2002). This suggested the existence of stable universal preferences (Walter

et al., 2020). In the literature, these partner preferences are often interchangeably referred to as mate preferences, ideal partner preferences, ideal standards or simply ideals.

1.2.1. *The Ideal Standards Model*

Ideal standards refer to the desires and expectations of a partner (Simpson et al., 2001). Research has found that these ideals can be grouped into three categories, namely characteristics related to 1) warmth and trustworthiness, 2) physical attractiveness and 3) status and resources (Fletcher et al., 1999). While more recent research replicated these dimensions, they also expanded on them with more categories describing an ideal partner such as dimensions related to confidence and humour (Gerlach et al., 2019), family orientation (Lam et al., 2016; Shackelford et al., 2005), intellect or dominance (Csajbók & Berkics, 2017; Schwarz & Hassebrauck, 2012). Yet, despite this universality of ideal partner preferences, people also show individual differences concerning the emphasis ascribed to each ideal partner category (Li et al., 2013; Watson et al., 2014; Zietsch et al., 2012).

Initially, I argued that from an evolutionary perspective it would be most adaptive if preferences guided the selection of a partner. The *Ideal Standards Model* (ISM) is based on this rationale, namely that partner selection is not random but guided by partner preferences (Fletcher et al., 1999). The ISM describes the selection of a partner as a continuous process in which the match between preferences and a partner's characteristics is intertwined with relationship decisions and outcomes (Fletcher et al., 1999; Fletcher et al., 2020). Further, the model assumes that preferences serve two functions. Those explain the way how preferences are related to decisions regarding the relationship (L. Campbell & Fletcher, 2015), henceforth referred to as relationship decisions.

First, partner preferences allow the evaluation of a (potential) partner and a relationship as preferences serve as a standard to which a partner or relationship can be compared to and, thus, influence a person's evaluation of the relationship. Indeed, research

has found a higher correspondence between preferences and a partner's characteristic to be associated with a higher relationship satisfaction and relationship quality for the person itself (e.g. Buyukcan-Tetik et al., 2017; L. Campbell et al., 2013; Fletcher et al., 2000) as well as for their partners (L. Campbell et al., 2001; L. Campbell et al., 2013). Furthermore, one study found that participants reported a lower relationship satisfaction when other individuals were more consistent with their preferences compared to their own partner. This was especially the case when participants had a higher mate value compared to their partners (Conroy-Beam et al., 2016).

Second, it is assumed that in case of a discrepancy between preferences and a partner's traits regulatory mechanisms are activated that reduce the discrepancy such as cognitive adjustments (e.g. lowering ideals or enhancing a partner's perception) or behavioural change (e.g. change the self or the partner) (L. Campbell & Fletcher, 2015). Support for this assumption comes from studies finding that individuals with a lower perceived consistency between ideals and the perception of a partner's traits show more attempts to change their partners (Overall et al., 2006). Studies have also found a smaller likelihood of separation if the consistency between preferences and a partner's traits is higher (Fletcher et al., 2000). However, results are not as straightforward as some studies cannot consistently replicate results (e.g. Eastwick et al., 2019; Eastwick & Neff, 2012; Lam et al., 2016). Those explain the way how preferences are related to decisions regarding the relationship (L. Campbell & Fletcher, 2015), henceforth referred to as relationship decisions.

1.2.2. Mixed Findings

An underlying assumption of the previously presented studies is that 1) the characteristics people report to prefer in a partner are linked to the characteristics of their actual partner and that 2) a match between their partner preferences and characteristics of

their partners is associated with relationship decisions (e.g. their relationship satisfaction). Both assumptions are referred to as the predictive validity of ideal partner preferences (L. Campbell & Stanton, 2014). However, many studies which investigated this assumption, examined already existing relationships. This study design suffers from the substantial limitation that a correspondence between preferences and a partner's traits can be the result of multiple alternative explanations. As the partner preferences participants had earlier to their current relationship are unknown, an alternative reason of a match could be that preferences are adjusted towards a partner, but did not influence initial relationship decisions.

Another possibility to investigate whether preferences are related to actual mate choices are speed-dating designs. In a speed-dating study, the initial stage of getting to know each other is operationalised with individuals interacting for a fixed period of time. After the initial interaction, participants indicate, for example, their attraction towards an interaction partner and whether they would like to meet the other person again. Researchers can then investigate whether partner preferences reported prior to the event, the so-called *stated preferences*, are related to *revealed preferences*, like the actual choice of meeting the interaction partner again. Yet, results of speed-dating studies are mixed.

On the one hand, speed-dating studies have found that initially reported preferences were not associated with actual choices. Instead, mainly physical attractiveness influenced participants' romantic interests (Asendorpf et al., 2011; Eastwick & Finkel, 2008; Joel et al., 2017; Todd et al., 2007). These findings could mean that participants are unaware of their preferences (Eastwick & Finkel, 2008). Because physical attractiveness was the main predictor of romantic interest, a *halo-effect (Thorndike, 1920)*, the attribution of positive impressions on other evaluations, could explain the pattern of results (for a discussion see Fletcher et al., 2014). A meta-analysis by Eastwick and colleagues (2014) and a recent literature review (Eastwick et al., 2019) corroborate the assumption that individuals are

unaware of their preferences that influence their partner selection: The authors have found that stated preferences were associated with choices in a hypothetical scenario, but not after live interactions as in the case of speed-dating studies. They propose that in live interactions, individuals might rather be guided by their feelings during these interactions instead of their previously reported preferences. This mounting literature casts doubt on the idea that reported preferences guide mate choices.

However, on the other hand, studies have found that, in addition to the main influence of physical attractiveness, preferences for age and ethnicity predicted actual choices (Kurzban & Weeden, 2005, 2007; D. Wood & Brumbaugh, 2009). Further, it has been found that women who reported a higher preference for masculinity also found a more masculine compared to a more feminine picture of the same person more attractive (DeBruine et al., 2006). One study even found that initially reported preferences for resources predicted romantic interest when more variance in a partner's characteristics was created (e.g. by recruiting a more diverse sample) (Li et al., 2013). In contrast to the evidence described before, these studies allow the interpretation that beyond physical attractiveness, other preferences influence romantic desire and attraction.

In summary, the continued research interest in what people desire in a partner stands in contrast to inconclusive findings on whether these preferences are actually influencing the selection of a partner at the initial stage of relationship formation and the assumption that a closer match between one's preferences and a partner's characteristics are in some form consequential.

1.3. Potential Reasons for these Mixed Findings

What could be the reasons for these ambiguous findings? Could it be that humans are unaware of their actual preferences that guide their relationship decisions? And that, if at all,

only the preference for the most obvious characteristics such as age and ethnicity influence their choices? Are partner preferences potentially unstable constructs which might explain the mixed findings? In the following, I am going to summarise possible explanations for these mixed findings grouped into four categories.

1.3.1. *Unrepresentative samples*

The first possible reason behind the mixed results could be due to unrepresentative samples. Most of the psychological research is based on student samples. This fact does not only constrain the representativeness of results, but also reduces the variation of important target-specific characteristics relevant in speed-dating studies (Fletcher et al., 2019). Consider, for example, the question whether an individual's stated preference for a partner with (prospective) resources is associated with choices made in a speed-dating design. Students might be too similar in this characteristic, which will limit the variance, and thus, will falsely lead to non-significant findings. This could explain the results of the speed-dating studies described before. Therefore, Li and colleagues (2013) manipulated the status of participant's interaction partners and found that with this higher variability in characteristics, now stated preferences were indeed associated with romantic interest.

Nonetheless, the limitation of using student samples applies to most, but not all of the speed-dating studies cited above. The studies by Asendorpf and colleagues (2011) and Todd and colleagues (2007) recruited participants from the broader population with a wider age range and still found that mainly physical attractiveness, but not stated preferences, predicted romantic interest. However, the latter might have been underpowered in order to find significant effects. Nevertheless, the study by Asendorpf and colleagues was adequately powered and used a community sample. Thus, it remains unclear whether unrepresentative samples are the sole reason for these mixed findings.

1.3.2. Investigated Time Frame

A second reason for the ambiguous findings could be the investigated time frame. Campbell and Stanton (2014) systematically examined the studies used in the meta-analysis by Eastwick and colleagues (2014). They realised that there was a gap in previous literature as studies either investigated the initial attraction towards a person (e.g. romantic interest at a speed-dating event) or already established relationships (e.g. influences on relationship satisfaction). What was missing were studies investigating the formation of relationships[3]. For this reason, the authors argued that based on the literature at that time, the conclusion drawn in the meta-analysis, namely that preferences have no influence on real-world scenarios, were too far reaching: it could only be concluded that stated preferences had a smaller effect at this initial stage, whereas preferences might still influence mate choices at later steps of relationship formation.

Based on this observation, two recent studies (L. Campbell et al., 2016; Gerlach et al., 2019) followed single participants over a time span of five months. Indeed, these studies found that initially stated preferences were linked to the characteristics of those who later became their romantic partners. However, at the initial point in time when the ideal partner preferences of single participants were assessed, they might have already known the partner with whom they had entered a relationship in the next five months. Put differently, the initially reported ideal partner preferences could have already been influenced by participants' perception of their later partners, and thus, explain the positive association of ideals and partner characteristics.

---

[3] Three studies tried to explore this process by investigating the formation of relationships after speed-dating events. However, the number of participants who actually met after the speed-dating events was too small to allow a reasonable examination.

Another major limitation of previous studies is that ideal partner preferences are assumed to be stable constructs (Fletcher et al., 1999). Their temporal stability over long periods of time, however, remains largely unexplored. Bleske-Rechek and colleagues (2015) have found a stability of $r = .35$ over a period of three years which is the longest retest interval investigated so far. However, since studies have found that ideals were adjusted towards a partner (Fletcher et al., 2000; Gerlach et al., 2019), partner preferences could also only be associated with characteristics of a partner over rather short periods of time. Over longer periods of time, the initially reported partner preferences might not be associated with characteristics of a partner and relationship decisions anymore.

Nevertheless, computer simulations are in line with the finding that preferences guide mate choices: A study has found that a preference-driven mate choice had more power to explain data of real-world couples as compared to a random mate choice (Conroy-Beam & Buss, 2016). In summary, the findings of studies that covered a wider time span, allow the interpretation that preferences could still play a role when selecting a partner. Nevertheless, this conclusion is still limited to a time span of five months, which highlights the need of studies investigating a longer period of time.

When investigating a longer time span, a sequential aspiration-level model (Miller & Todd, 1998) could provide a theoretical framework on how preferences might be integrated into mating decisions. At the initial stage of getting to know each other, preferences for characteristics which are easy to observe are most influential on partner selection. Since these characteristics can also be perceived more reliably as compared to, for example, the reliability of a person, it seems perfectly reasonable to base relationship decisions on accurately perceived characteristics first, in this case on physical attractiveness (Fletcher et al., 2014; Miller & Todd, 1998). At later stages of relationship formation, during which further characteristics could be perceived more reliably, other preferences could then come

into play. As proposed by the sequential aspiration-level model, this preference integration could follow the rule that in order for a person to be still deemed as an eligible partner for a next stage of relationship formation, a certain characteristic needs to exceed a certain minimum threshold (i.e. a person only takes others into account as a potential partner if they exceed a minimum level of physical attractiveness). Yet, there are also alternative theories on how these preferences could be integrated with first evidence suggesting that a partner's characteristics could be compared to one's ideals instead of only exceeding a minimum threshold (Conroy-Beam, 2021). However, this field on how preferences are exactly integrated into mate choices is still in its infancy (see Conroy-Beam, 2021 or Brandner et al., 2020, for the first studies really testing models on how preferences could be integrated into mate choices).

1.3.3. *Study Design as a Confounding Factor*

As mentioned above, Eastwick and colleagues (2019) interpreted their literature review insofar that preferences had no real-life impact on mate choices but only influenced hypothetical scenarios. However, Fletcher and colleagues (2020) introduced a second interpretation. They argued that studies reporting effects versus those with non-significant findings might be confounded by the study's design. More precisely, studies that reported significant effects of preferences investigated hypothetical scenarios in which variables were manipulated. For example, in the study by DeBruine and colleagues (2006) pictures of the same person where shown to a rater, one of which being edited to look more masculine and the other one being made to look more feminine. Initially, each rater reported his or her preference for masculinity, and then had to select the picture which he or she perceived to be more attractive. Results showed that a rater's initially reported preference for masculinity was associated with his or her attractiveness ratings of a picture. In contrast, studies investigating face-to-face contexts (e.g. speed-dating studies), that did not use an experimental approach,

often reported non-significant findings. The choice of the study's design ultimately influenced the resulting data structure and thereby the statistical power to detect an effect. In studies investigating face-to-face contexts, the investigated interaction between two continuous variables (ideal partner preference * characteristics) requires a high statistical power to detect an effect. In contrast, experiments, in which the target's characteristics are manipulated, test for main effects, and therefore require lower statistical power (Rohrer & Arslan, 2021). In addition to this difference in analytical approaches, the experimental manipulation of characteristics could have made them more salient to participants, which in turn, increased the influence of preferences and thereby the chance to produce statistically significant effects (for more details see Fletcher et al., 2020). The assumption of Fletcher and colleagues is underpinned by one exception: The study by Li and colleagues (2013) investigated a face-to-face context, but used experimental manipulation and found that stated preferences were, indeed, related to mate choices. The authors recruited two groups of participants: university students as well as participants with a low social status. Hence, the authors manipulated the participant's status. They found that individuals who reported a higher importance of social status prior to the event, also preferred individuals with a higher status at the speed-dating event (e.g. reported higher romantic interest for these individuals).

1.3.4. Debated Statistical Approaches

Finally, a reason for the mixed findings could lie in the diverse statistical approaches that were used to analyse data regarding the same research question. A recent debate on the "best practices" to investigate whether preferences influence relationship decisions arose (see Eastwick et al., 2019; Fletcher et al., 2020). An important part of the debate was whether an association between ideals and a partner's characteristics could be interpreted as evidence for a preference-driven mate choice. One side claimed that a significant association could be

owed to multiple other reasons[4] (Eastwick et al., 2019) and was therefore not informative since it could not be interpreted as evidence for a preference driven mate choice. The other side argued that the association between ideals and partner characteristics was indeed informative. They argued that a significant association gives room to the possibility of a preference driven mate choice even if there were other reasons that could have led to the same pattern of results. Only if a non-significant association is found, then this null finding can be interpreted as initial evidence that preference do not influence mate choices (Fletcher et al., 2020).

Another part of the debate concentrated on how the data has previously been analysed. As aforementioned, Eastwick et al. (2019) raised concerns about correlating ideals and partner characteristics in order to investigate the influence of preferences on mate choices. Instead, they suggested that a better test would be to investigate whether a match between ideals and partner characteristics was in any form consequential with regard to relationship outcomes. However, the authors pointed out that most of the previous studies following this statistical approach suffered from one limitation, namely, that they did not control for the *Normative-Desirability Confound* (NDC) (D. Wood & Furr, 2016). NDC refers to the finding that random profiles of people are similar not because of their actual similarity, but because of their shared similarity with an average person, the so-called normative profile. At the same time, this normativity is confounded with the desirability of characteristics, meaning that more normative characteristics are also socially more desirable. The authors argued that this confounding could have inflated the similarity between ideals and partner characteristics, which might have led to spurious findings that a closer match is indeed associated with better relationship outcomes. They support their claim with the few studies that controlled for the NDC and revealed ambigous evidence whether a match is

---

[4] Alternative explanations are described in more detail in the discussion (see 5.2. Statistical Analyses).

associated with a higher relationship quality (Eastwick et al., 2019). In conclusion, with different researchers suggesting different approaches, the debate about the best statistical approach is still ongoing and no consensus has been reached, except from the agreement that several statistical approaches should be pursued.

1.4. Theoretical Conclusion and Aim of this Dissertation

So far, studies found that mainly physical attractiveness influenced romantic desire in speed-dating studies, thereby casting doubt on the role of preferences on relationship decisions. These findings stand in stark contrast to a majority of studies finding that humans can clearly state what they are seeking for in an ideal partner and the assumption that a match between preferences and a partner's characteristics is related to relationship outcomes. It was only until recently that studies began to close this gap by investigating the role of preferences at later stages of relationship formation. Although these studies provide some evidence that preferences could guide the selection of a partner, conclusions can only be drawn over rather short periods of time. It remains largely unclear what happens to the influence of preferences on relationship decisions over longer time spans.

This dissertation aims to follow-up on the shortcomings of previous literature and to re-evaluate the influence of preferences on relationship decisions. We strived to remove concerns raised on statistical approaches, the representativeness of samples and investigate a much wider time span with multiple study designs. In order to do so, my colleagues and I made use of two approaches. In the first step, we focused on initial attraction. We especially zoomed in on the role of intelligence as it belongs to the most highly valued characteristics of a partner (e.g. Buss & Barnes, 1986; Li et al., 2002). In a second approach, we investigated preferences and their role on mating decisions more broadly in a longitudinal design. Although studies never investigated the stability of preferences over longer periods than three

years, they are assumed to be stable cognitive constructs (Fletcher et al., 1999). But since research has found that ideals are adjusted towards a partner (Gerlach et al., 2019), it is unclear whether partner preferences can be related to characteristics of a partner over longer periods of time. We set out to test whether the claim of stable ideal partner preferences is justified and investigated their stability over a period of 13 years. In the third and final approach, we investigated whether preferences reported 13 years ago were associated with the perceived characteristics of later found partners and whether a closer match was associated with relationship outcomes such as satisfaction and commitment.

2.   Summary of Manuscript 1

In the first manuscript, my co-authors and I investigated the role of partner preferences on initial mating decisions. We were specifically interested in the role of intelligence on initial attraction. Why did we specifically investigate the role of intelligence? First, as cross-culturally, intelligence belongs to the most highly valued characteristics in an ideal partner (Buss & Barnes, 1986; Walter et al., 2020), it seems to be an important criterion in human mate choice. Second, we wanted to test an influential theory of evolutionary psychology proposing that the preference for an intelligent partner could have evolved due to sexual selection (Miller, 2000b, 2000a). This assumption is based on the theoretical background that having an intelligent partner has several direct benefits as intelligence is linked to many positive outcomes in life like health, socio-economic status, or income (Deary, 2012). In the light of the strong link between harmful mutations and a well-functioning brain, intelligence could have also evolved as an indicator of genetic quality, referred to as a fitness indicator (Hawrylycz et al., 2012; Klasios, 2013). Thus, having an intelligent partner could also provide indirect benefits (Miller, 2000b, 2000a). However, up to date, most previous studies have only investigated the role of perceived intelligence on

attraction, but not the role of objectively measured intelligence (Karbowski et al., 2016; e.g. Lee et al., 2014; but see Prokosch et al., 2009). Consequently, we investigated whether individuals with a higher objectively measured intelligence had a higher mate appeal in two complementing research designs (a highly controlled rating study and a speed-dating study) in which the weaknesses of one study were buffered by the other. Additionally, we investigated whether funniness was associated with measured intelligence as it was assumed to be a perceptible cue of intelligence (Miller, 2000a).

For our rating study, 88 target men were photographed and video-taped performing several tasks (referred to as cues) after a comprehensive measure of their intelligence. These cues were selected in order to gradually increase the information on the target's intelligence. More precisely, we recorded cues mainly revealing men's physical attractiveness (facial and full-body photograph), a cue controlling for their vocal attractiveness (a video of men reading out loud all vowels) and cues which allow an accurate perception of intelligence (Borkenau et al., 2004) and display men's humour (a video of men performing a pantomime task, a video in which men read aloud newspaper headlines, and a video of men telling an anecdote). 179 women then rated either men's funniness, intelligence or mate appeal based on a selection of these cues. In our second study, 763 individuals participated in a speed-dating event after we assessed their verbal intelligence. At each event, participants interacted for three minutes with two to five partners. After each interaction, participants rated their partners on several attributes including their mate appeal, funniness and intelligence.

In both studies, measured intelligence was positively associated with perceived intelligence, suggesting that intelligence was perceptible even after short interactions. However, we found that only perceived, but not measured intelligence predicted mate appeal. Rather, physical attractiveness was the main predictor of mate appeal. Funniness was only associated with the perception of being intelligent, but not with measured intelligence.

Consequently, these findings do not support the assumption that intelligence influences initial attraction, casting doubt on the sexual selection theory of intelligence as proposed by Miller (Miller, 2000b, 2000a). Additionally, our pattern of results seems to be partly due to a halo effect, as physically attractive individuals were perceived more positively across all other domains. The relevance of a halo effect for initial attraction is further supported by results of a very recent study reporting similar results (Hofer et al., 2021). All things considered, these findings cast doubt on the idea that intelligence evolved through sexual selection. Integrating these results into the broader framework, we replicated findings that physical attractiveness is most decisive for initial attraction. However, intelligence might still play a role at a later stage of relationship formation.

3.   Summary of Manuscript 2

The aim of the second manuscript focusses more broadly on improving the understanding of the concept of ideal partner preferences in general. My co-authors and I investigated whether ideal partner preferences were stable over a period of 13 years. We additionally explored potential changes and participant's insights into these changes. Stable ideal partner preferences are a necessary condition if we believe that preferences influence relationship decisions over longer periods of time. Previous research has shown that preferences seem to be stable over three years, although their stability substantially reduced from $r = .85$ after three weeks (Fletcher et al., 1999) to $r = .35$ after three years (Bleske-Rechek & Ryan, 2015). As proposed in the ISM, a reason for this reduced stability could be that a discrepancy between ideals and a partner's characteristics could activate cognitive processes (e.g. adjusting one's preference) in order to reduce this mismatch (L. Campbell & Fletcher, 2015). This assumption is supported by previous research finding that ideals are altered in order to lower the discrepancy towards a partner's characteristics (Fletcher et al.,

2000; Gerlach et al., 2019). This means that preferences are potentially a constant subject to change. Consequently, initial preferences might only guide relationship decisions over rather short periods of time. In contrast, over longer periods of time, altered preferences, but not initial preferences, might be related to mate choices. However, this question remains largely unanswered as, to our knowledge, no study investigated the stability of ideal partner preferences exceeding a time span of three years.

To answer the question whether ideal partner preferences are stable over a much longer period of time, we asked 204 participants to rate the importance of 58 characteristics in an ideal partner at two time points in their life, once in 2006 and a second time 13 years later in 2019. In 2019, they also indicated whether they perceived to have changed their ideal partner preferences across the time. We decided to assess participants perceived preference change because it has been hypothesised that if participants had an accurate insight, this could eliminate the need for costly longitudinal studies (Sprecher et al., 2018). Indeed, our results indicated that preferences had a substantial stability over 13 years. Although speaking of small effect sizes ($d = .15 - .22$), we found that participants reported an increased preference for a partner with a higher status and resources and a partner who is trustworthy and warm. Participants' preferences for an attractive and vital partner rather decreased over the course of 13 years. Changes in preferences were related to participants' age and whether they were planning on having or had already started a family. For example, the rank-order stability of preferences was significantly lower for individuals who had the intention to have a family or already started having a family compared to individuals without (the intention to have) children. However, participant's insight into these changes varied, emphasising the need of longitudinal studies and suggesting that the perception of change cannot be used as a substitute of the assessment of actual changes. Overall, these findings laid the foundations for

our third manuscript, since we learned that initial preferences could potentially guide

relationship decisions across much longer time spans than just three years.

4.   Summary of Manuscript 3

In the third manuscript, we investigated the same sample as in the second manuscript.

Here, we tested whether the ideal partner preferences reported in 2006, when participants

were still single, were associated with characteristics of romantic partners in the following 13

years and whether relationship outcomes were associated with a closer match between partner

preferences and a partner's perceived characteristics.

For this purpose, we asked participants to fill out an event history calendar (EHC) in

2019, in which they listed all relationships they have had since 2006 that lasted longer than

six months and also their current relationship independent of its length. Since participants had

to have been in at least one relationship over the course of 13 years, our sample size for

investigating our research question was slightly reduced to 178 participants who have had

322 relationships since 2006. Participants then described each relationship in more detail (e.g.

they rated their relationship satisfaction). In case that their relationship was still ongoing,

participants were also asked to rate their commitment to their current relationship. With

regard to ex-partners, participants were asked to report who had ended the relationship.

Participants also rated each partner regarding 25 characteristics, whereby the content of these

characteristics was congruent to participants' reported ideal partner preferences in 2006. We

aimed to validate participants' ratings of their partners with their partners' self-reports and

with participants' friends or family members who were asked to rate participants' partners.

Unfortunately, only few partners and peers took this opportunity. Although participants'

ratings of their partners were substantially correlated with their partners' self-reports and peer

reports of their partners' characteristics, we did not use this data for subsequent analyses due

to the small sample size. Hence, our analyses are based on target perceptions of a partner's characteristics only.

We predicted perceived partner traits with participant's ideals as reported in 2006. We found that participants' initial preferences were associated with their perception of their partners, allowing the interpretation that preferences do indeed guide mate choices. We then investigated whether a match of preferences and the perception of a partner has consequences on a relationship. To this end, we analysed whether a closer match between preferences and perceived partner characteristics was associated with a higher relationship satisfaction and the length of a relationship. For still ongoing relationships, we investigated the influence of a match on participants' commitment and in cases of ex-partners, we were interested in whether a match was associated with who ended the relationship. In light of the current debate on best practices in analytic approaches (see Eastwick et al., 2019; Fletcher et al., 2020), we adopted several statistical approaches to test the robustness of our findings. Accordingly, we predicted relationship outcomes with 1) the distance and 2) the correlation between ideals and perceived partner traits. Across all analyses, we found that participants reported a higher commitment to their relationship with a closer match between their preferences and perception of their partners' characteristics. Results varied for all other relationship outcomes. Whereas effects were in the predicted direction (a closer match being associated with better relationship outcomes) for less strict analyses, effects diminished in more conservative analyses correcting for biases in perceptions. Hence, it remains unclear whether matching ideal preferences and perceived partner characteristics affect other relationship outcomes except for commitment. While we discussed several possibilities for this unclear pattern, we concluded that the best statistical approach might not have been found yet.

5.   General Discussion

During the past years, the role of preferences in human relationship decisions has been controversially debated. Many influential theories proposed that preferences guide the selection of a partner. Despite findings supporting these theories, there are contrasting findings that question whether preferences are in any way influential and suggest that the selection of a partner is rather random. In order to advance the scientific discourse, the aim of this dissertation was to re-evaluate the influence of preferences on relationship decisions using two distinct approaches. In the beginning, I raised four issues which might explain ambiguity in previous findings. In what follows, I am going to elaborate on how we addressed each concern in our manuscripts and what our findings tell us about the influence of preferences on relationship decisions.

5.1. Unrepresentative samples

One point of critique referred to the fact that the majority of previous studies relied on student samples which could result in diminished variance in important characteristics that might have led to difficulties in detecting meaningful effects (Fletcher et al., 2019).

We partly addressed this limitation in our first manuscript. This manuscript consists of two studies, a speed-dating and a rating study. In the speed-dating study, we still relied on a student sample and unfortunately have only limited ability to investigate the variance of the outcome variable. However, we particularly targeted to overcome this problem in the rating study by recruiting target participants across the whole city centre assuming that we would end up with a sample of the broader population and a wider range of intelligence scores. Considering the broad range of intelligence scores present in this more diverse sample, we achieved this goal. Hence, we cannot preclude that the aforementioned limitation might still apply to the speed-dating study, but we can rule out that it applies to the rating study.

Nevertheless, in both we found that intelligence has no effect on initial attraction. These results suggest that in the specific case of intelligence on initial attraction, a limited variance in the outcome variable does not seem to be the reason for non-significant findings.

Regarding our sample collected for the second and third manuscript, we were able to fully address the issue of unrepresentative samples. Our sample consisted of participants recruited by Asendorpf and colleagues (2011), the only large community sample so far (as mentioned above). We re-contacted these participants of the original speed-dating study in 2019. Since this sample was recruited from the broader population, it deviates from the traditional student sample and has a much wider age range (Manuscript 2: range = 31 - 66; Manuscript 3: range = 31 - 66). Although we faced a large dropout rate of 41% until 2019, our sample size is still comparable to other studies in this field. Because we now found that partner preferences were related to perceived partner characteristics and relationship commitment, a reduced variance in the outcome variable could explain non-significant findings of previous studies. But concerning the influence of intelligence on initial attraction, this limitation in previous studies does not seem to be the sole reason for null findings, and thus, overall this suggests that limited target variability cannot be the mere reason for non-significant findings of previous studies.

5.2. Statistical analyses

Recently, the best statistical approach to analyse whether partner preferences guide relationship decisions has been debated. Two main concerns were raised: First, the sheer association between partner preferences and partner characteristics is not an unequivocal test of the hypothesis whether preferences guide the selection of a partner, since alternative explanations could have produced the same pattern of results. For example, an alternative explanation of why ideals are associated with partner characteristics could be the

confounding with potential third-variables (Fletcher et al., 2020). Second, analyses should correct for the normative desirability of the investigated characteristics. As this point of critique is most relevant to the analyses of Manuscript 3, I am going to discuss the issues based on our third manuscript.

Regarding the first issue, one such confounding third variable could be related to the social environment a person lives in. To provide an example: if people in one's social environment are highly educated, one might believe that education is highly important in an ideal partner. Moreover, this person has a higher probability of finding a highly educated partner in this environment (Gerlach et al., 2019). This possible confounding factor was named *passive ideal change* (Eastwick et al., 2019). Another alternative explanation could be an unclear direction of causality which Eastwick and colleagues (2019) defined as motivated ideal change (e.g. because one's partner has a good education, one therefore believes that education is important in an ideal partner). With our third manuscript, we cannot assess how far participants' ideals are shaped by their social environment. Social homogamy clearly seems to affect where we meet our partners. However, we assessed participants' ideals when they were still single before they had selected their partners. This means that at least a motivated ideal change is unlikely. Thus, we believe that our findings are an initial indicator that preferences could be the driving force in mate choice. Nevertheless, as we cannot rule out every alternative explanation of why we found an association between preferences and perceived partner characteristics, we also investigated whether a match is in any form consequential. More precisely, we investigated whether a closer match between ideals and perceived partner characteristics is associated with relationship outcomes, such as relationship satisfaction or commitment.

In order to address the second concern regarding the normative desirability confound (NDC), we initially aimed to validate perceived partner characteristics with two additional

sources. We tried to collect partners' self-reports and the perceptions of peers on each partners' characteristics. With this extension of our study, we intended to receive more objective ratings and potentially address the NDC. Unfortunately, a small number of peers and partners participated in this second part of the study, eliminating reasonable analyses based on this data. That is why we had to rely on alternative ways to control for biases in participants' perceptions of their partners. Therefore, we applied several approaches, including those correcting for the NDC, in order to investigate whether a closer match between preferences and perceived partner characteristics was related to better relationship outcomes. Across all analyses, a closer match was associated with a higher relationship commitment. Findings were, however, more ambiguous for all other outcomes. In the less strict analyses (not or only partly correcting for the NDC) a match of ideals and perceived partner characteristics was associated with relationship outcomes such as satisfaction or relationship length. Nonetheless, results were not significant in the more conservative analyses. In the manuscript, we discussed the strengths and weaknesses of each approach in more detail. However, we came to the conclusion that the approach correcting for the NDC might not be applicable in the context of partner selection. The reason is of technical nature: The NDC is determined by calculating the average perception of a partner (the mean of each perceived partner characteristic across the sample). Next, the NDC is eliminated by subtracting this average profile from each partner perception (the average perception of each characteristic is subtracted from each corresponding characteristic of a partner perception), resulting in their unique profile. This is perfectly reasonable for studies in which the sample is representative. However, in studies on partner selection, as was the case here, this procedure of determining the normative profile results in the average profile of a person who has been selected as a partner. But in fact, our analyses suggested that this profile is different from the profile of people who do not enter a relationship. Consequently, correcting for the

NDC in studies on partner selection might be biased due to an unrepresentative correction. Instead, future studies could calculate the normative desirability of each characteristic in more representative samples (including individuals who are single over longer periods of time). To sum up, we addressed all concerns raised in the debate about the most appropriate statistical approach, but believe that it still stands to reason whether an optimal approach has been found yet.

5.3. Investigated Time Frame

Researchers highlighted the need for studies that do not only investigate the initial stage of relationship formation or already existing relationships, but also the time span in between (L. Campbell & Stanton, 2014). Although we investigated the role of intelligence only on initial attraction in Manuscript 1, we specifically addressed this issue in Manuscripts 2 and 3. Two recent studies (L. Campbell et al., 2016; Gerlach et al., 2019) tracked single participants over five months and found that initial preferences were associated with characteristics of later partners. Because preferences seem to be adjusted towards a partner (Gerlach et al., 2019), it remained unclear whether preferences would still play a role over longer periods than just five months. It seems as if we came one step closer to closing the gap of studies investigating the role of preferences over longer time spans since our manuscripts cover a time period starting from initial attraction up to 13 years later.

5.4. Study Design as a Confounding Factor

The last point of critique that has to be addressed was that only studies using experimental manipulation, but not studies in face-to-face contexts, found preferences to be related to relationship decisions. One important concern that needs to be addressed here is that experimental manipulation might have led to a greater power to detect an effect due to

the resulting data structure and increasing participants' awareness of their preferences

(Fletcher et al., 2020).

In our first manuscript, we investigated the influence of intelligence on mate appeal using a face-to-face context (our speed-dating study) and an experimental manipulation (our rating study). In both studies, we did not find that intelligence influenced mate appeal, regardless of the study's design. In the speed-dating study that used live interactions, we investigated the main effect of intelligence on mate appeal in a large sample reducing the possibility of the study being underpowered. In the rating study, we even manipulated the amount of information on the target's intelligence and could still not find that intelligence influenced mate appeal. Consequently, our results do not show that the concern is justified. However, both studies only investigated the role of intelligence on initial attraction. Hence, they do not allow to draw any conclusions on the role of other preferences and over longer periods of time. Additionally, it has to be noted that we assumed that every participant had a similar high preference for a more intelligent partner, which might not be applicable as individual differences in the preference for intelligence could exist.

In Manuscript 3, we investigated the influence of individual partner preferences on actual relationships. Consequently, we did not conduct an experimental manipulation, but rather investigated real-life scenarios. In contrast to other studies investigating face-to-face contexts, we found that preferences are related to the perception of a partner's characteristics and to participants' commitment to their relationship. What is the difference between our study and other studies investigating face-to-face contexts that were unable to find an effect? Most of these studies used speed-dating designs, whereas ours investigated a timeframe of 13 years. Although differences in study-designs could explain the null-findings of previous studies, the investigated time frame could have played the decisive role.

5.5. Do Preferences Influence Relationship Decisions?

Summarising previous concerns in studies investigating the role of preferences on relationship decisions, we named four central problems in previous research and addressed each in our studies in four important ways: We 1) recruited more representative samples, 2) approached the question with two diverging study designs reducing a possible confounding, 3) investigated not only initial attraction or already established relationships, but also the process in between, and finally, 4) tackled our research questions with several, more sophisticated statistical approaches. Embedding the results of the dissertation into the broader context, allows to re-evaluate the question whether preferences influence attraction and partner selection. Supported by Manuscript 1, physical attractiveness seems to be most decisive for initial attraction and intelligence might play an inferior role. Furthermore, results of Manuscript 2 indicate that participants have substantially stable ideal partner preferences. Stability in preferences allows that preferences guide relationship decisions across long time frames. Finally, the results of Manuscript 3, namely that 1) partner preferences are associated with the perceived characteristics of partners found later, and that 2) a closer match between preferences and partner perceptions is related to commitment, suggest that preferences are, indeed, consequential and guide relationship decisions. Thus, our results support the predictive validity of ideal partner preferences. Figure 1 summarises the findings of each manuscript and illustrates at which stages of relationship formation this dissertation investigated the influence of preferences on relationship decisions.

Figure 1

Illustration at Which Stages of Relationship Formation the Dissertation Investigated Whether Preferences are Related to Relationship Decisions
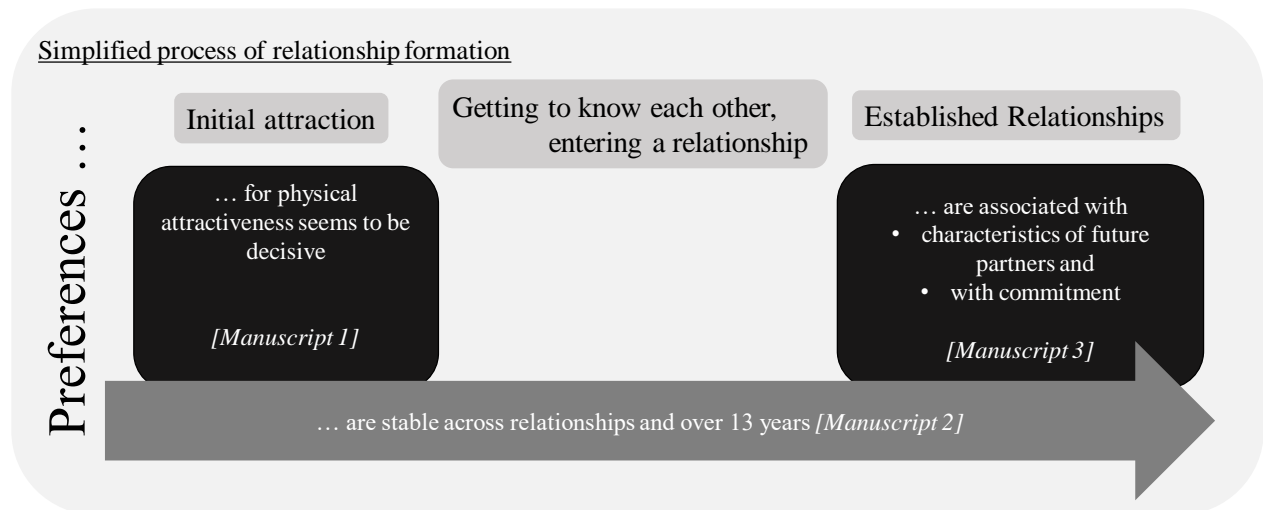
Results across the manuscripts can be embedded within a sequential process model that defines different stages of relationship formation at which preferences influence relationship decisions. Moreover, this model is able to explain the mixed results of previous research. Characteristics that are easy to observe, such as physical attractiveness, influence initial attraction. Supporting this notion, Fletcher and colleagues (2014) have found that physical attractiveness was more reliably perceived than status and resources or warmth and trustworthiness of a person. The authors emphasised that integrating only reliably perceived characteristics into relationship decisions is adaptive because it reduces the costs of making a wrong decision. Although we found that intelligence was accurately perceived after short interactions, the perception might still not be reliable enough in order to be used at initial encounters. An alternative explanation could also be that although participants' perception of intelligence was sufficiently accurate, it did not influence their rating of a target's mate appeal because they rather evaluated each target in terms of a short-term mate, in which intelligence might play a minor role compared to physical attractiveness. Consequently, actual intelligence does not seem to play a role for initial attraction. This would also explain

why speed-dating studies could not find that stated preferences were related to actual choices made at the event, since characteristics other than physical attractiveness cannot be perceived reliable or are not important to individuals for initial attraction. However, over time, participants will, most likely, be able to perceive other characteristics such as the trustworthiness or reliability, as well as the intelligence of a potential partner more accurately. Across these later stages of relationship formation, their preferences for these characteristics could then guide their relationship decisions. This could explain our finding that preferences were related to perceived partner characteristics and that a match between both influenced participants' relationship commitment.

Nevertheless, so far, it remains unclear at what exact point in time this might occur in the process of relationship formation. Ideally, future study should extend on our study design: Such a study should follow a community sample over a period of several years from individuals being single until many years later and include several measurement points (e.g. panel data). At each assessment, characteristics should be assessed of individuals who were selected but also rejected as a potential partner, thus optimally a person's whole social network should be investigated. Ideally, these characteristics are assessed not only with a person's perception of another individual, but also with objective ratings of another person's characteristics (e.g. by recruiting at least one other person within a person's network).

5.6. Can Alternative Models Explain our Pattern of Results?

Based on the ISM, we assumed that participants would compare a (potential) partner to their ideals and base their mate choice upon their ideal preferences. However, alternative models exist that describe how preferences could be integrated into mate choices (for an overview see Conroy-Beam, 2021). Could these models better explain our pattern of results?

Miller and Todd (1998) proposed a sequential aspiration-level model, meaning that a person's characteristic has to exceed a certain threshold in a desired trait before this person is deemed an eligible partner. In support of this sequential aspiration-level model, Lee and colleagues (2014) showed that women had to exceed a minimum level of physical attractiveness, before the perception of her intelligence influenced men's ratings of the woman's attractiveness of her dating website's profile. However, one study also found that an IQ of 120 was rated as most attractive, whereas individuals with an IQ exceeding 135 were not rated as attractive anymore (Gignac et al., 2018). This finding implies that characteristics of potential partners do not need to exceed a certain threshold, but need to lie in a certain range in order to influence relationship decisions. Applying this model to our findings for intelligence preferences, it is possible that intelligence does not affect initial attraction because everyone was literate and thus, exceeded this minimum threshold of a preferred intelligence level. In this case, we cannot rule out that this alternative model would explain our pattern of results.

Further research investigated how such a threshold or ideal could be determined. One possibility is that it is based on participants' own mate values, which participants learn to gauge throughout adolescence (Penke et al., 2008). Starting from this self-assessment, humans seek a partner who is similar or even exceeds their own mate value (Bruch & Newman, 2018; Figueredo et al., 2006; Watson et al., 2014). Since we only assessed participants' but not their partners' mate value, we cannot test whether this model could explain our results. Making matters even more complex, another possibility is, that ideals could interact with each other. For example, Jensen-Campbell and colleagues (1995) have found an interaction between agreeableness and dominance: Women rated more dominant men only as desirable partners, if men had a high compared to a low level of agreeableness. Buyukcen-Tektik and colleagues (2017) proposed that one pronounced partner characteristic

could compensate for another lower characteristic. It is possible, that even multiple

mechanisms are intertwined: For example, for intelligence, it could be that a potential partner

needs to exceed a certain threshold but for trustworthiness, the potential partner has to be

close to one's ideals. With regard to physical attractiveness, a highly attractive person might

compensate for a lack of resources.

In conclusion, I cannot disentangle which exact processes are at work when

preferences are integrated into relationship decisions. However, in line with our findings,

recent studies support the idea that a partner's characteristics are compared to ideals (Conroy-

Beam, 2018; Conroy-Beam & Buss, 2016). Nevertheless, multiple mechanisms could also be

at work which would explain our pattern of results as well. This strongly highlights the need

for future research on how preferences are integrated into mate choices.

5.7. Limitations

Despite the many strengths of the studies described in this dissertation (e.g. the wide

time span or more diverse samples), there were also some limitations. Our inferences drawn

from Manuscript 1 are limited to the role of intelligence at initial encounters. Additionally,

we assumed that all participants would highly prefer a more intelligent partner. Hence, we

cannot say what role intelligence plays after initial attraction. Moreover, it might be possible

that we did not address a potential influence of inter-individual differences in participants'

preferences for intelligence on relationship decisions, which we hope future research will

shed light on.

Another limitation consists of the assessment of a partner's characteristics, relevant in

Manuscript 3. Despite our effort to collect more objective ratings of a partner's

characteristics, we could only rely on our participants' perceptions of their partners. These

perceptions are potentially biased (Eastwick et al., 2019; but see Fletcher et al., 2014).

Although we approached potential biases with sophisticated analyses, these analyses might not have been optimal.

Finally, in Manuscript 2 and 3, we only had two measurement time points, meaning that we could not investigate the processes in-between initial attraction up to the point where a relationship is established. As a result, we do not know how preferences developed over the 13 years in between. Moreover, we only have information about people who were selected as a partner, but not about those individuals who might have been rejected as a potential partner for a committed relationship. A promising line for future studies that might be able to address this issue is investigating the whole social network of a person. Such studies may shed light on when and how preferences change over the course of relationships and show which preferences influence relationship decisions across different stages of relationship formation. Eastwick and colleagues (2021) started to adopt such an approach and asked students every 3 weeks over the course of around 6 months to name and rate potential partners in their social network. However, their study is limited to the early development of a relationship in a sample of participants in their first semester at university, and thus, cannot disentangle the exact process on how preferences are integrated into relationship decisions. Thus, there is a need for future research to reach a more profound knowledge on how and when preferences are integrated into relationship decisions.

6. Conclusion

In this dissertation, I am able to re-evaluate the influence of partner preferences on relationship decisions based on three complementing manuscripts: First, my co-authors and I investigated initial attraction and second, we investigated an extensive time span of 13 years. This investigation fills a formerly dire gap in the literature as, to my knowledge, no study investigated such a long time span so far. Moreover, I described shortcomings of previous

research which might explain the mixed results on whether preferences affect relationship decisions. I showed that our manuscripts were able to overcome these burdens. Thereby, the results of all three manuscripts were able to show that preferences do play a role in relationship decisions. Moreover, a sequential model could explain of how preferences influence relationship decisions. However, it still stands to reason on how exactly preferences are integrated, highlighting the need for future research.

We have not only learned that there is some truth in commercials for dating websites, but that these findings are much more important on a more profound level. Previously, I described that romantic relationships have a major importance on our health and well-being. For example marriage is associated with a higher life expectancy, people are in general happier and less depressed (Holt-Lunstad et al., 2010; Holt-Lunstad, 2018; Perelli-Harris et al., 2019). At the same time unhappy marriages are related to the reversed pattern, like a higher chance of an earlier death (Sbarra et al., 2011). We have found that a higher match between preferences is related to a higher commitment in relationships and – although evidence is weaker – it could also be related to relationship satisfaction. This means with more research on what makes a happy relationship, we could design more sophisticated interventions for relationship counselling. More broadly, this could buffer the negative effects of an unhappy relationship for not only an individual but also our society.

7.   References

Asendorpf, J. B., Penke, L., & Back, M. D. (2011). From Dating to Mating and Relating: Predictors of Initial and Long–Term Outcomes of Speed–Dating in a Community Sample. *European Journal of Personality*, *25*(1), 16–30. https://doi.org/10.1002/per.768

Bleske-Rechek, A., & Ryan, D. E. (2015). Continuity and change in emerging adults' mate preferences and mating orientations. *Personality and Individual Differences*, *72*, 90–95. https://doi.org/10.1016/j.paid.2014.08.033

Borkenau, P., Mauer, N., Riemann, R., Spinath, F. M., & Angleitner, A. (2004). Thin Slices of Behavior as Cues of Personality and Intelligence. *Journal of Personality and Social Psychology*, *86*(4), 599–614. https://doi.org/10.1037/0022-3514.86.4.599

Brandner, J. L., Brase, G. L., & Huxman, S. A. (2020). "Weighting" to find the right person: compensatory trait integrating versus alternative models to assess mate value. *Evolution and Human Behavior*, *41*(4), 284–292. https://doi.org/10.1016/j.evolhumbehav.2020.05.001

Bruch, E. E., & Newman, M. (2018). Aspirational pursuit of mates in online dating markets. *Science Advances*, *4*(2), eaap9815. https://doi.org/10.1126/sciadv.aap9815

Buss, D. M. (1989). Sex differences in human mate preferences: Evolutionary hypotheses tested in 37 cultures. *Behavioral and Brain Sciences*, *12*(1), 1–14. https://doi.org/10.1017/S0140525X00023992

Buss, D. M., & Barnes, M. (1986). Preferences in human mate selection. *Journal of Personality and Social Psychology*, *50*(3), 559–570. https://doi.org/10.1037/0022-3514.50.3.559

Buss, D. M., & Schmitt, D. P. (1993). Sexual Strategies Theory: An evolutionary perspective on human mating. *Psychological Review*, *100*(2), 204–232. https://doi.org/10.1037/0033-295x.100.2.204

Buss, D. M., & Schmitt, D. P. (2019). Mate Preferences and Their Behavioral Manifestations. *Annual Review of Psychology*, *70*(1), 77–110. https://doi.org/10.1146/annurev-psych-010418-103408

Buyukcan-Tetik, A., Campbell, L., Finkenauer, C., Karremans, J. C., & Kappen, G. (2017). Ideal Standards, Acceptance, and Relationship Satisfaction: Latitudes of Differential Effects. *Frontiers in Psychology*, *8*, 1691. https://doi.org/10.3389/fpsyg.2017.01691

Byrne, D. (1961). Interpersonal attraction and attitude similarity. *The Journal of Abnormal and Social Psychology*, *62*(3), 713–715. https://doi.org/10.1037/h0044721

Byrne, D., & Blaylock, B. (1963). Similarity and assumed similarity of attitudes between husbands and wives. *The Journal of Abnormal and Social Psychology*, *67*(6), 636–640. https://doi.org/10.1037/h0045531

Campbell, L., Chin, K., & Stanton, S. C. E. (2016). Initial Evidence that Individuals Form New Relationships with Partners that More Closely Match their Ideal Preferences. *Collabra*, *2*(1). https://doi.org/10.1525/collabra.24

Campbell, L., & Fletcher, G. J. O. (2015). Romantic relationships, ideal standards, and mate selection. *Current Opinion in Psychology*, *1*, 97–100. https://doi.org/10.1016/j.copsyc.2015.01.007

Campbell, L., Overall, N. C., Rubin, H., & Lackenbauer, S. D. (2013). Inferring a partner's ideal discrepancies: Accuracy, projection, and the communicative role of interpersonal behavior, *105*(2), 217–233. https://doi.org/10.1037/a0033009

Campbell, L., Simpson, J. A., Kashy, D. A., & Fletcher, G. J. O. (2001). Ideal Standards, the Self, and Flexibility of Ideals in Close Relationships. *Personality and Social Psychology Bulletin*, *27*(4), 447–462. https://doi.org/10.1177/0146167201274006

Campbell, L., & Stanton, S. C. E. (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*(9), 485–494. https://doi.org/10.1111/spc3.12126

Christensen, H. T. (1947). Student Views on Mate Selection. *Marriage and Family Living*, *9*(4), 85–88. https://doi.org/10.2307/347505

Conroy-Beam, D. (2018). Euclidean Mate Value and Power of Choice on the Mating Market. *Personality and Social Psychology Bulletin*, *44*(2), 252–264. https://doi.org/10.1177/0146167217739262

Conroy-Beam, D. (2021). Couple Simulation: A Novel Approach for Evaluating Models of Human Mate Choice. *Personality and Social Psychology Review*, *25*(3), 191–228. https://doi.org/10.1177/1088868320971258

Conroy-Beam, D., & Buss, D. M. (2016). How Are Mate Preferences Linked with Actual Mate Selection? Tests of Mate Preference Integration Algorithms Using Computer Simulations and Actual Mating Couples. *PLoS ONE*, *11*(6), e0156078. https://doi.org/10.1371/journal.pone.0156078

Conroy-Beam, D., Goetz, C. D., & Buss, D. M. (2016). What predicts romantic relationship satisfaction and mate retention intensity: mate preference fulfillment or mate value discrepancies? *Evolution and Human Behavior*, *37*(6), 440–448. https://doi.org/10.1016/j.evolhumbehav.2016.04.003

Csajbók, Z., & Berkics, M. (2017). Factor, factor, on the whole, who's the best fitting of all? Factors of mate preferences in a large sample. *Personality and Individual Differences*, *114*, 92–102. https://doi.org/10.1016/j.paid.2017.03.044

Darwin, C. (1871). *The Descent of Man and Selection in Relation to Sex*. Murray.

Deary, I. J. (2012). Intelligence. *Annual Review of Psychology*, *63*(1), 453–482. https://doi.org/10.1146/annurev-psych-120710-100353

DeBruine, L. M., Jones, B. C., Little, A. C., Boothroyd, L. G., Perrett, D. I., Penton-Voak, I. S., Cooper, P. A., Penke, L., Feinberg, D. R., & Tiddeman, B. P. (2006). Correlated preferences for facial masculinity and ideal or actual partner's masculinity. *Proceedings of the Royal Society B: Biological Sciences*, *273*(1592), 1355–1360. https://doi.org/10.1098/rspb.2005.3445

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*, *92*(2), 245–264. https://doi.org/10.1037/0022-3514.94.2.245

Eastwick, P. W., Finkel, E. J., & Simpson, J. A. (2019). Best Practices for Testing the Predictive Validity of Ideal Partner Preference-Matching. *Personality and Social Psychology Bulletin*, *45*(2), 167–181. https://doi.org/10.1177/0146167218780689

Eastwick, P. W., Joel, S., Molden, D. C., Finkel, E. J., & Carswell, K. L. (2021). *Predicting Romantic Interest during Early Relationship Development: A Preregistered Investigation using Machine Learning* [Manuscript submitted for publication]. University of California, Davis.

Eastwick, P. W., Luchies, L. B., Finkel, E. J., & Hunt, L. L. (2014). The predictive validity of ideal partner preferences: A review and meta-analysis. *Psychological Bulletin*, *140*(3), 623–665. https://doi.org/10.1037/a0032432

Gangestad, S. W., & Scheyd, G. J. (2005). The Evolution of Human Physical Attractiveness. *Annual Review of Anthropology*, *34*(1), 523–548. https://doi.org/10.1146/annurev.anthro.33.070203.143733

Gerlach, T. M., Arslan, R. C., Schultze, T., Reinhard, S. K., & Penke, L. (2019). Predictive validity and adjustment of ideal partner preferences across the transition into romantic relationships. *Journal of Personality and Social Psychology*, *116*(2), 313–330. https://doi.org/10.1037/pspp0000170

Gignac, G. E., Darbyshire, J., & Ooi, M. (2018). Some people are attracted sexually to intelligence: A psychometric evaluation of sapiosexuality. *Intelligence*, *66*, 98–111. https://doi.org/10.1016/j.intell.2017.11.009

Gonzaga, G. C., Carter, S., & Buckwalter, J. G. (2010). Assortative mating, convergence, and satisfaction in married couples. *Personal Relationships*, *17*(4), 634–644. https://doi.org/10.1111/j.1475-6811.2010.01309.x

Hawrylycz, M. J., Lein, E. S., Guillozet-Bongaarts, A. L., Shen, E. H., Ng, L., Miller, J. A., van de Lagemaat, L. N., Smith, K. A., Ebbert, A., Riley, Z. L., Abajian, C., Beckmann, C. F., Bernard, A., Bertagnolli, D., Boe, A. F., Cartagena, P. M., Chakravarty, M. M., Chapin, M., Chong, J., . . . Jones, A. R. (2012). An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature*, *489*(7416), 391–399. https://doi.org/10.1038/nature11405

Hofer, G., Burkart, R., Langmann, L., & Neubauer, A. C. (2021). What you see is what you want to get: Perceived abilities outperform objective test performance in predicting mate appeal in speed dating. *Journal of Research in Personality*, *93*, 104113. https://doi.org/10.1016/j.jrp.2021.104113

Holt-Lunstad, J. (2018). Why Social Relationships Are Important for Physical Health: A Systems Approach to Understanding and Modifying Risk and Protection. *Annual Review of Psychology*, *69*(1), 437–458. https://doi.org/10.1146/annurev-psych-122216-011902

Holt-Lunstad, J., Smith, T. B., & Layton, J. B. (2010). Social Relationships and Mortality Risk: A Meta-analytic Review. *PLOS Medicine*, *7*(7), e1000316. https://doi.org/10.1371/journal.pmed.1000316

Jensen-Campbell, L. A., Graziano, W. G., & West, S. G. (1995). Dominance, prosocial orientation, and female preferences: Do nice guys really finish last? *Journal of Personality and Social Psychology*, *68*(3), 427–440. https://doi.org/10.1037/0022-3514.68.3.427

Jia, H., & Lubetkin, E. I. (2020). Life expectancy and active life expectancy by marital status among older U.S. adults: Results from the U.S. Medicare Health Outcome Survey (HOS). *SSM - Population Health*, *12*, 100642. https://doi.org/10.1016/j.ssmph.2020.100642

Joel, S., Eastwick, P. W., & Finkel, E. J. (2017). Is Romantic Desire Predictable? Machine Learning Applied to Initial Romantic Attraction. *Psychological Science*, *28*(10), 1478–1489. https://doi.org/10.1177/0956797617714580

Kalmijn, M. (1998). Intermarriage and Homogamy: Causes, Patterns, Trends. *Annual Review of Sociology*, *24*(1), 395–421. https://doi.org/10.1146/annurev.soc.24.1.395

Karbowski, A., Deja, D., & Zawisza, M. (2016). Perceived female intelligence as economic bad in partner choice. *Personality and Individual Differences*, *102*, 217–222. https://doi.org/10.1016/j.paid.2016.07.006

Klasios, J. (2013). Cognitive Traits as Sexually Selected Fitness Indicators. *Review of General Psychology*, *17*(4), 428–442. https://doi.org/10.1037/a0034391

Kurzban, R., & Weeden, J. (2005). HurryDate: Mate preferences in action. *Evolution and Human Behavior*, *26*(3), 227–244. https://doi.org/10.1016/j.evolhumbehav.2004.08.012

Kurzban, R., & Weeden, J. (2007). Do advertised preferences predict the behavior of speed

daters? *Personal Relationships*, *14*(4), 623–632. https://doi.org/10.1111/j.1475-

6811.2007.00175.x

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*(6), 703–722.

https://doi.org/10.1177/0146167216637842

Lee, A. J., Dubbs, S. L., Hippel, W. von, Brooks, R. C., & Zietsch, B. P. (2014). A

multivariate approach to human mate preferences. *Evolution and Human Behavior*, *35*(3),

193–203. https://doi.org/10.1016/j.evolhumbehav.2014.01.003

Li, N. P., Bailey, J. M., Kenrick, D. T., & Linsenmeier, J. A. W. (2002). The necessities and

luxuries of mate preferences: Testing the tradeoffs. *Journal of Personality and Social

Psychology*, *82*(6), 947–955. https://doi.org/10.1037/0022-3514.82.6.947

Li, N. P., Yong, J. C., Tov, W., Sng, O., Fletcher, G. J. O., Valentine, K. A., Jiang, Y. F., &

Balliet, D. (2013). Mate preferences do predict attraction and choices in the early stages of

mate selection. *Journal of Personality and Social Psychology*, *105*(5), 757–776.

https://doi.org/10.1037/a0033777

Luo, S. (2017). Assortative mating and couple similarity: Patterns, mechanisms, and

consequences. *Social and Personality Psychology Compass*, *11*(8), e12337.

https://doi.org/10.1111/spc3.12337

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 and Social Psychology*,

65(1), 56–68. https://doi.org/10.1037/0022-3514.65.1.56

Miller, G. F. (2000a). *The mating mind: How sexual choice shaped the evolution of human

nature*. Doubleday.

Miller, G. F. (2000b). Mental Traits as Fitness Indicators - Expanding Evolutionary

Psychology's Adaptationism. In D. LeCroy & P. Moller (Eds.), *Evolutionary perspectives*

on human reproductive behavior (pp. 62–74). Annals of the New York Academy of

Sciences. https://doi.org/10.1111/j.1749-6632.2000.tb06616.x

Miller, G. F., & Todd, P. M. (1998). Mate choice turns cognitive. *Trends in Cognitive*

Sciences, *2*(5), 190–198. https://doi.org/10.1016/S1364-6613(98)01169-3

Newcomb, T. M. (1956). The prediction of interpersonal attraction. *American Psychologist*,

11(11), 575–586. https://doi.org/10.1037/h0046141

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(4), 662–685. https://doi.org/10.1037/0022-3514.91.4.662

Penke, L., Todd, P. M., Lenton, A. P., & Fasolo, B. (2008). How self-assessments can guide

human mating decisions. In *Mating intelligence: Sex, relationships, and the mind's*

reproductive system (pp. 37–75). Lawrence Erlbaum Associates Publishers.

Perelli-Harris, B., Hoherz, S., Lappegård, T., & Evans, A. (2019). Mind the "Happiness"

Gap: The Relationship Between Cohabitation, Marriage, and Subjective Well-being in the

United Kingdom, Australia, Germany, and Norway. *Demography*, *56*(4), 1219–1246.

https://doi.org/10.1007/s13524-019-00792-4

Pietromonaco, P. R., & Collins, N. L. (2017). Interpersonal mechanisms linking close

relationships to health. *American Psychologist*, *72*(6), 531–542.

https://doi.org/10.1037/amp0000129

Prokosch, M. D., Coss, R. G., Scheib, J. E., & Blozis, S. A. (2009). Intelligence and mate

choice: intelligent men are always appealing. *Evolution and Human Behavior*, *30*(1), 11–

20. https://doi.org/10.1016/j.evolhumbehav.2008.07.004

Rohrer, J. M., & Arslan, R. C. (2021). Precise Answers to Vague Questions: Issues With Interactions. *Advances in Methods and Practices in Psychological Science*, *4*(2), 25152459211007368. https://doi.org/10.1177/25152459211007368

Sbarra, D. A., Law, R. W., & Portley, R. M. (2011). Divorce and Death: A Meta-Analysis and Research Agenda for Clinical, Social, and Health Psychology. *Perspectives on Psychological Science*, *6*(5), 454–474. https://doi.org/10.1177/1745691611414724

Schwarz, S., & Hassebrauck, M. (2012). Sex and Age Differences in Mate-Selection Preferences. *Human Nature*, *23*(4), 447–466. https://doi.org/10.1007/s12110-012-9152-x

Shackelford, T. K., Schmitt, D. P., & Buss, D. M. (2005). Universal dimensions of human mate preferences. *Personality and Individual Differences*, *39*(2), 447–458. https://doi.org/10.1016/j.paid.2005.01.023

Simpson, J. A., Fletcher, G. J. O., & Campbell, L. (2001). The Structure and Function of Ideal Standards in Close Relationships. In G. J. O. Fletcher & M. S. Clark (Eds.), *Blackwell Handbook of Social Psychology: Interpersonal Processes* (pp. 86–106). Blackwell Publishers.

Sprecher, S., Econie, A., & Treger, S. (2018). Mate preferences in emerging adulthood and beyond: Age variations in mate preferences and beliefs about change in mate preferences. *Journal of Social and Personal Relationships*, *36*(10), 3139–3158. https://doi.org/10.1177/0265407518816880

Symons, D. (1979). *The evolution of human sexuality*. Oxford University Press.

Thorndike, E. L. (1920). A Constant Error in Psychological Ratings. *Journal of Applied Psychology*, *4*(1), 25–29. https://doi.org/10.1037/h0071663

Todd, P. M., Penke, L., Fasolo, B., & Lenton, A. P. (2007). Different cognitive processes underlie human mate choices and mate preferences. *Proceedings of the National Academy of Sciences*, *104*(38), 15011–15016. https://doi.org/10.1073/pnas.0705290104

Trivers, R. L. (1972). Parental investment and sexual selection. In B. G. Campbell (Ed.), *Sexual Selection and the descent of man* (pp. 136–179). Aldine.

Walter, K. V., Conroy-Beam, D., Buss, D. M., Asao, K., Sorokowska, A., Sorokowski, P., Aavik, T., Akello, G., Alhabahba, M. M., Alm, C., Amjad, N., Anjum, A., Atama, C. S., Atamtürk Duyar, D., Ayebare, R., Batres, C., Bendixen, M., Bensafia, A., Bizumic, B., . . . Zupančič, M. (2020). Sex Differences in Mate Preferences Across 45 Countries: A Large-Scale Replication. *Psychological Science*, *31*(4), 408–423. https://doi.org/10.1177/0956797620904154

Watson, D., Beer, A., & McDade-Montez, E. (2014). The Role of Active Assortment in Spousal Similarity. *Journal of Personality*, *82*(2), 116–129. https://doi.org/10.1111/jopy.12039

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*(5), 1029–1068. https://doi.org/10.1111/j.0022-3506.2004.00289.x

Wood, D., & Brumbaugh, C. C. (2009). Using revealed mate preferences to evaluate market force and differential preference explanations for mate selection. *Journal of Personality and Social Psychology*, *96*(6), 1226–1244. https://doi.org/10.1037/a0015300

Wood, D., & Furr, R. M. (2016). The Correlates of Similarity Estimates Are Often Misleadingly Positive: The Nature and Scope of the Problem, and Some Solutions. *Personality and Social Psychology Review*, *20*(2), 79–99. https://doi.org/10.1177/1088868315581119

Wood, W., & Eagly, A. H. (2012). Chapter two - Biosocial Construction of Sex Differences and Similarities in Behavior. In James M. Olson & Mark P. Zanna (Eds.), *Advances in Experimental Social Psychology* (Vol. 46, pp. 55–123). Academic Press. https://doi.org/10.1016/B978-0-12-394281-4.00002-7

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, *9*(2, Pt.2), 1–27. https://doi.org/10.1037/h0025848

Zhang, L., Lee, A. J., DeBruine, L. M., & Jones, B. C. (2019). Are Sex Differences in Preferences for Physical Attractiveness and Good Earning Capacity in Potential Mates Smaller in Countries With Greater Gender Equality? *Evolutionary Psychology*, *17*(2), 1474704919852921. https://doi.org/10.1177/1474704919852921

Zietsch, B. P., Verweij, K. J. H., & Burri, A. V. (2012). Heritability of Preferences for Multiple Cues of Mate Quality in Humans. *Evolution*, *66*(6), 1762–1772. https://doi.org/10.1111/j.1558-5646.2011.01546.x

Appendix A.

Manuscript 1

*Julie Driebe and Morgan Sidari share the first authorship

**Lars Penke, Brendan Zietsch and Ruben Arslan share the last authorship

Title:

Intelligence can be detected but is not found attractive in videos and live interactions

Driebe, Julie C.*[a,1]

Sidari, Morgan J.*[a,2]

Dufner, Michael[3]

von der Heiden, Juliane M.[4]

Bürkner, Paul C.[5]

Penke, Lars[b, 1,6]

Zietsch, Brendan P.[b,2]

Arslan, Ruben C.[b,7]

[a] share first authorship, [b] share last authorship

* corresponding author for study 1: julie.driebe@uni-goettingen.de, for study 2: morgan.sidari@uqconnect.edu.au

[1] Biological Personality Psychology, University of Goettingen, Germany

[2] Centre for Psychology and Evolution, School of Psychology, University of Queensland, Australia

[3] Personality Psychology and Psychological Assessment, Witten/Herdecke University, Germany

[4] Personality Psychology and Psychological Assessment, University of Mainz, Germany

[5] Computer Science, Aalto University, Finnland

[6] Leibniz ScienceCampus Primate Cognition, Goettingen, Germany

[7] Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin, Germany

Abstract

Self-reported mate preferences suggest intelligence is valued across cultures, consistent with the idea that human intelligence evolved as a sexually selected trait. The validity of self-reports has been questioned though, so it remains unclear whether objectively assessed intelligence is indeed attractive. In Study 1, 88 target men had their intelligence measured and based on short video clips were rated on intelligence, funniness, physical attractiveness and mate appeal by 179 women. In Study 2 ($N = 729$), participants took part in 2 to 5 speed-dating sessions in which their intelligence was measured and they rated each other's intelligence, funniness, and mate appeal. Measured intelligence did not predict increased mate appeal in either study, whereas perceived intelligence and funniness did. More intelligent people were perceived as more intelligent, but not as funnier. Results suggest that intelligence is not important for initial attraction, which raises doubts concerning the sexual selection theory of intelligence.

Keywords: intelligence, mate choice, sexual selection

Humans' extraordinary intelligence is an important aspect that distinguishes us from all other animals. However, the evolutionary forces that gave rise to this peculiar feature are not well understood. Our intelligence seems to go far beyond what is required for mere survival, as it enables us to compose music, create art and literature, and to engage in humorous wordplay. Such activities do not have clear survival benefits, and indeed the human brain's energy demands are enormous relative to the other organs of the human body and the brains of other animals (Mink et al., 1981). One theory is that our surplus of intelligence has emerged through intersexual selection (Miller, 2000b, 2000a), which results from individual differences in attractiveness to the opposite-sex (Darwin, 1871). Specifically, Miller (2000a) proposed that intelligence serves as a fitness indicator to potential mates. As 84% of human genes are expressed in the brain, developing a healthy, optimally functioning brain requires an individual to be relatively free from harmful mutations (Hawrylycz et al., 2012; Klasios, 2013; Miller, 2000b, 2000a). For this reason, intelligence, or displays that require intelligence, such as humour, may signal genetic quality to potential romantic partners.

If human intelligence and humour evolved via romantic and sexual choices across multitudes of generations, this legacy should be reflected in our romantic and sexual preferences today (Miller, 2000b, 2000a; Puts, 2010). Accordingly, research has found that intelligence and humour are reported as among the most desirable traits in a hypothetical ideal partner (Bressler & Balshine, 2006; Buss et al., 1990; Li et al., 2002; Sprecher & Regan, 2002; Wilbur & Campbell, 2011); however, other studies have shown that these ideal partner preferences do not closely correspond to mate preferences revealed in attraction to real individuals (Eastwick et al., 2011). To test whether intelligence is truly predictive of mate appeal, research should not rely on self-reported partner preferences, but rather have participants rate the mate appeal of individuals who also had their intelligence tested objectively. Prokosch et al. (2009) conducted such a study, providing some evidence that women were more attracted to men (in videos performing verbal and physical tasks) who scored higher on a measure of intelligence. However, only 15 men were involved in the study, so the evidence should be regarded as preliminary. Other studies have connected measured intelligence and humour production in writing tasks (Greengross & Miller, 2011; Howrigan & MacDonald, 2008), but no study has tested whether measured intelligence relates to humour as it is used in live interactions, which is the relevant case in terms of the evolutionary question. In all, the attractiveness of intelligence and its relation to interpersonal

humour remain open questions that are key to the viability of the sexual selection theory of these traits.

The Current Study

Here, we conducted two studies to investigate the accuracy of intelligence judgements based on short sequences of behaviour (Ambady & Rosenthal, 1992) as well as the impact of intelligence on mate appeal and perceptions of funniness. In study 1, we used highly controlled conditions (i.e. short video sequences of participants), comprehensive intelligence measures, a large target sample size, and a repeated measures design that assessed women's judgements multiple times as the information on targets' intelligence increased. The purpose of this repeated measures design is that by gradually presenting different cues with increasing intelligence information above cues on only physical attractiveness, we can isolate the effect of intelligence on mate appeal (see Fig. 1). In addition, different samples of women rated either intelligence, funniness, physical attractiveness or mate appeal to reduce transfer effects and shared response tendencies. These design features allow us to determine how mate appeal of targets changes with more information about their intelligence and funniness while, importantly, allowing us to control for potential halo effects. According to Miller's hypothesis, the preference for intelligence should be stronger among female, as compared to male, perceivers (Buss et al., 1990; Sprecher & Regan, 2002). Hence, testing women's preferences is a powerful test of the hypothesis.

In study 2, we adopted a more ecologically valid speed-dating design whereby participants' verbal intelligence was measured and they provided ratings on each other's intelligence, funniness and mate appeal after a 3-minute meeting.

Study 1. For intelligence to play a focal role in human mate choice, it needs to be perceived somewhat accurately. First, we predict that women's intelligence ratings for male targets, based on short sequences of behaviour (e.g. reading newspaper headlines aloud), will be positively correlated with targets' psychometrically measured intelligence.

Second, we investigate the influence of funniness, a proposed more perceivable display of intelligence, on sexual mate appeal. We hypothesise that perceived funniness is associated with measured intelligence and that men's perceived funniness will predict their rated sexual mate appeal above and beyond the effect of their intelligence. Further, we hypothesise that perceived intelligence predicts rated sexual mate appeal.

Third, in line with Miller's (2000a) hypothesis, we predict that men's measured intelligence will be significantly positively correlated with women's ratings of men's sexual mate appeal. And fourth, we predict that the increase in men's mate appeal after adding

additional cues related to intelligence (i.e. reading newspaper headlines aloud; making experimenter laugh) will depend on men's intelligence, such that the sexual mate appeal increase will be greater for more intelligent men.

Study 2. For study 2, the hypotheses follow a similar rationale. First, we predict that psychometrically measured intelligence will be positively correlated with speed-dating partners' perception of intelligence. Second, we predict that measured intelligence will be positively correlated with speed-dating partners' ratings of mate appeal and funniness. Third, we predict speed-dating partners' ratings of intelligence and ratings of funniness will be positively correlated with their ratings of mate appeal for the same target.

Method

Study 1

Parts of study 1 were preregistered (https://osf.io/rs3tg); however, during the course of the project we realised that some specifications were insufficient and we opted for more appropriate analyses. For transparency, we have provided a table in our appendix (S1) which highlights the deviations from our preregistration and details their respective rationales.

Data collection for study 1 was completed in three steps: an online questionnaire and follow-up lab-based session with male participants (stimuli), and several lab-based sessions with female participants (raters). All participants provided written consent and were informed about the study's aim after participation. Studies like ours are exempt from IRB according to German regulations.

Participants

Male targets. An online survey titled 'Person Perception' was used to screen participants for inclusion in our lab-based study. Participants were recruited with posters in the city centre (e.g. train stations, gyms, job centres) and the X university campus. Of the 347 participants that commenced the survey, 118 males finished[5]. All of these 118 men over the age of 18 years were then recruited to participate in our lab-based study. Final participants were 88 males with ages ranging from 19 to 31 years ($M = 24.22$, $SD = 2.81$). Participation was incentivised through a small payment (10€) and personalised feedback on their personality. The sample varied in educational attainment, ranging from university degrees (26%), high school degrees (67%), vocational baccalaureate diploma (5%), to secondary

---

[5] Of these 347 participants, 169 only clicked on our survey. Another 35 participants were female. Hence, 169 men started filling out our online survey with 118 finishing our online study participation.

school leaving certificates (2%). The vast majority of the sample was heterosexual (97%), with one homosexual and two bisexual participants. The majority of men were single (61%) and the remainder were currently in a romantic relationship (39%).

Female raters. Participants were recruited through various online channels (e.g. Facebook, a local student participant pool) and posters on campus. Of the 203 participants that responded, 24 were excluded on the basis of either being male (14), technical difficulties (9), or previous participation (1). We also excluded ratings in which women reported acquaintance with the male target, leaving a final number of 39,003 ratings (3% dropout) from 179 women with ages ranging from 18 to 36 years ($M = 21.84$, $SD = 3.22$). Participation was incentivised through a coupon lottery and course credit for those recruited at the university. The vast majority of the sample was heterosexual (93%), with one homosexual participant (1%), and 11 bisexual (6%) participants; 55% were in relationships and 45% were single.

Participants were distributed across six rating studies (described in greater detail in S2) with the sample size breakdown as follows: study 1.1 ($n = 19$, ratings = 1657), 1.2 ($n = 16$, ratings = 1368), 1.3 ($n = 30$, ratings = 2620), 1.4 ($n = 25$, ratings = 10,485), 1.5 ($n = 30$, ratings = 12,739), and 1.6 ($n = 59$, ratings = 10,134). Demographics for individual groups are reported in the supplementary materials (S2).

Materials and Procedure

Male targets. Participants completed an online questionnaire implemented via the survey framework formr.org (Arslan et al., 2019). The questionnaire included basic demographic items (e.g. age, gender, sexual orientation, and educational attainment), along with more extensive measures related to intelligence (extended German version of the International Cognitive Ability Resource ICAR; Condon & Revelle, 2014), and personality (irrelevant to the current study). Each subsequent laboratory session, which yielded the stimuli for study 1, lasted approximately one hour and was conducted by the same two female experimenters to standardise experimenter effects across participants and induce potential effects of female presence on male self-display behaviour (Ronay & Hippel, 2010).

At the beginning of the session, additional assessments of men's measured intelligence were applied, namely the Deary-Liewald Reaction Time Task (DLRT; Deary et al., 2011), the Multiple Choice Vocabulary Test (MWT-B; Lehrl, 2005), and the knowledge scale from the Berliner Test zur Erfassung Fluider und Kristalliner Intelligenz (BEFKI GC-K; Schipolowski et al., 2013). Men were then photographed and videotaped performing several tasks (see Fig. 1). First, a facial photograph (cue 1) and second a full body photograph (cue 2)

of men standing on a marked spot to standardise lighting and focal distance was taken. Men received no instructions for posture and facial expression. Third, we videotaped men reading vowels out loud (cue 3). Each vowel was displayed onscreen for two seconds each to standardise reading speed. Fourth, the men were videotaped while reading five newspaper headlines from German newspapers aloud as this task is strongly related to an accurate intelligence perception (Borkenau et al., 2004). In order for them to be intellectually challenging, we selected headlines containing foreign words or describing complex facts (e.g. 'Compensation payments lead US diocese into bankruptcy.'). Fifth, we videotaped men pantomiming the word 'Zahnrad' (mechanical gear) which we used as a warm-up and the word 'Bankverbindung' (bank details) (cue 5). Last, men were asked to make the experimenter laugh within a 30 seconds time limit by telling an anecdote or joke (cue 6); they were given five minutes to prepare for this task prior to video recording. Full HD cameras (resulting in a resolution of 1920 x 1080 pixels) were used for all recordings and clips were created with the program Mangold VideoSyncPro IP Version 1.7.0.22.

Fig. 1.
Overview of study 1 stimuli.
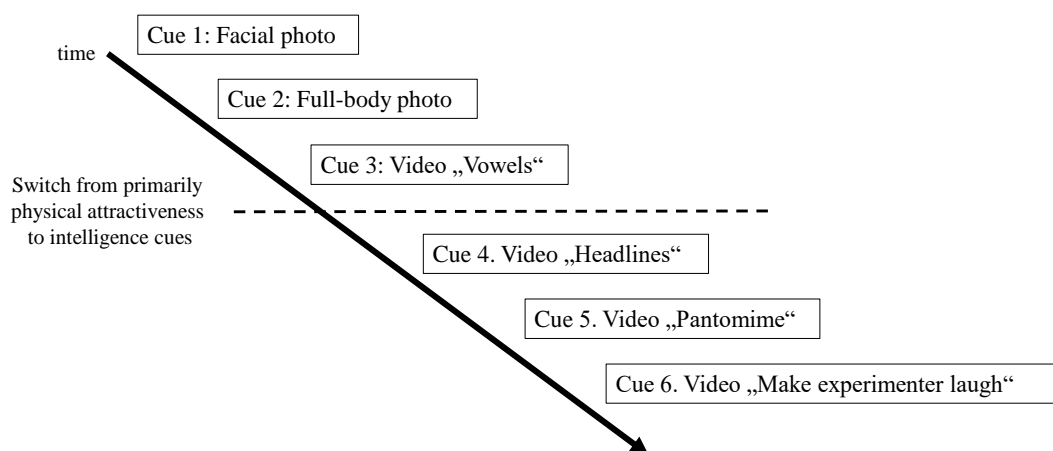
Female raters. Female raters participated in one of six computer-based rating studies (referred to herein as rating study 1.1 - 1.6) based on slightly different sets of stimuli. For all rating studies, the session began with a short demographic questionnaire, including age, gender, educational attainment, relationship status, and sexual orientation. Rating study 1.1 assessed a baseline of men's physical attractiveness, 1.2 assessed perceived intelligence and

funniness, and 1.3 assessed men's attractiveness as a short-term mate and long-term mate. Rating study 1.4 assessed changes in men's short-term mate attractiveness when shifting from physical attractiveness information (cues 1-3) to additional cues related to men's intelligence (cue 4 and cue 6). Rating studies 1.5 and 1.6 were replications of rating study 1.4 with small methodological improvements. Stimuli were randomised into two blocks: after watching the first block, participants were able to take a 15-minute break to reduce test fatigue. In rating study 1.6, women only rated a randomly drawn half of our target sample (44 men) to further reduce test fatigue; in all other studies, all 88 men were rated. Studies 1.1 to 1.4 were programmed using the Software PsychoPy2 Experiment Builder (v1.80.06) (Peirce, 2007); however, a software update of PsychoPy crashed experiment 1.5, therefore, we ran study 1.6 and the majority of study 1.5 on the experimental framework Alfred (Treffenstaedt & Wiemann, 2018).

Rating study 1.1. Participants rated the target's physical attractiveness after being shown two photographs (cue 1: facial photograph; cue 2: full body photograph). The item ('How attractive do you find this man?') was rated on a 7-point Likert-scale ranging from 1 (not attractive at all) to 7 (very attractive).

Rating study 1.2. Participants rated targets' intelligence and funniness after watching three video sequences of each target (cue 4-6). The item (e.g. 'He is intelligent', 'He is humorous') was rated on a 5-point Likert-scale ranging from 1 (strongly disagree) to 5 (strongly agree).

Rating study 1.3. Participants watched the same three videos as in rating study 1.2 (i.e. cue 4-6); however, they were instead asked to evaluate men's short term- and long-term mate attractiveness. The items ('How well can you imagine having a sexual affair with this man?' and 'How well can you imagine a long-term relationship with this man?') were rated on a 7-point Likert-scale ranging from 1 (not at all) to 7 (very well).

Rating study 1.4. Participants were provided with definitions of short-term mate ('brief sexual encounters') and long-term mate ('serious, committed relationships') prior to ratings. Participants provided ratings five times: first after they saw facial photographs (cue 1), then after seeing full body photographs (cue 2), then after seeing each of three additional videos (cue 3, 4, and 6). Cue 5 was not presented in order to reduce test fatigue. Each time the item ('Please rate the following recording of this man considering his short-term and long-term mate attractiveness') was rated on two separate response scales ranging from 1 (not attractive) to 100 (very attractive).

Rating study 1.5. The procedure for rating study 1.5 was almost identical to rating study 1.4; however, participants were now instructed to evaluate men's short-term and long-term mate attractiveness independently of their own relationship status. That is, women were asked to provide ratings from the perspective of a single woman even if they were partnered. Additionally, women saw a preview of all 88 facial photographs of the target men prior to making any responses. These modifications were made because the ratings in the first study were extremely low (mean of 19 on a scale from 0 to 100), suggesting a floor effect. By previewing the full range of men in the study, we hoped that women would not reserve their highest attractiveness rating in the expectation that a more attractive man would appear. For the preview, each man's picture was displayed for two seconds in a randomised order. As a final attempt to improve discrimination between targets, we also explicitly pointed out the whole range of the scale to participants.

Rating study 1.6. The procedure for rating study 1.6 slightly improved upon rating study 1.5 with an aim of reducing potential fatigue effects. In this study, twice the number of female raters rated half of the targets (44 of 88). Additionally, women saw men's facial and full body photographs (cue 1 and 2, respectively) and made their first rating based on both photos. The items were phrased identically to rating studies 1.4 and 1.5; however, the scale now ranged from -50 (repulsive) to +50 (attractive). The slider was preset to the scale's midpoint (0).

Statistical Analyses

All our analyses were run using R 3.6.0 (R Core Team, 2019).

Male stimuli. Targets' measured intelligence, extracted as a *g* factor, is the first unrotated factor of a principal component analysis of the eight intelligence tests used in study 1.

Accuracy of intelligence perception. For each male target, we aggregated all women's ratings of men's intelligence to calculate the *aggregated perceiver accuracy*. We correlated men's actual intelligence with this aggregated perceived intelligence to investigate the accuracy of intelligence perception. Additionally, we fitted a structural equation model in *lavaan* v0.6-4 (Rosseel, 2012) modelling *g* as a hierarchical latent variable to correct for measurement error and clustering standard errors by target to estimate the semi-latent single rater accuracy.

Preference for Intelligence. To test whether intelligence adds a unique contribution to men's long-term and short-term mate attractiveness, we used Bayesian multilevel linear models calculated in Stan (Carpenter et al., 2017) with the *brms* package v 2.10.0 (Bürkner,

2017) with weakly informative priors. To validate our analyses, we additionally fitted models in *lme4* v1.1-21 (Bates et al., 2015). As ratings resulted from three different studies (rating studies 1.4-1.6), we included an interaction between study and each cue, allowing for varying influences of cues on long-term mate/short-term mate ratings in each study. Because the studies grouped cues differently, the cue variable had four levels that were entered dummy-coded: face/body photo, voice, newspaper headlines and make experimenter laugh, with the voice recording set as the reference category. Of main interest, we specified population-level interactions between the cues and intelligence. These were adjusted for by specifying interactions between cues and physical attractiveness. We specified varying intercepts for targets and raters. Additionally, we allowed the effect of the cue dummy variables to differ between targets and the interaction between cues and traits to differ by rater. Finally, we let an interaction between cue and study and varying intercepts for raters and targets predict the residual standard deviation in the regression in a location-scale model to account for the fact that the rating scale might be used differently across studies and participants.

        Preference for funniness. To assess the influence of funniness incremental to the influence of measured intelligence on mate appeal, we regressed men's *g* factor and ratings of their funniness onto their mate appeal. We used the packages *sandwich* v2.5-1 (Berger et al., 2017; Zeileis, 2004) and *lmtest* v0.9-37 (Zeileis & Hothorn, 2002) to correct our standard errors as ratings of men's physical attractiveness, funniness and mate appeal were clustered in three different sets of female raters.

        Robustness Checks. We stated in the preregistration that we would only recruit heterosexual raters and targets, so we repeated all of our analyses excluding participants who indicated that they were not heterosexual. We also stated in the preregistration that we would use aggregated ratings instead of women's individual ratings for a given trait. Those aggregations were planned for physical attractiveness, short-term mate attractiveness, long-term mate attractiveness, perceived intelligence and perceived funniness. We conducted these analyses as a robustness check.

Study 2

Participants

        Participants were 729 (379 female) first year psychology students with ages ranging from 16.92 to 36.08 years (women: $M = 19.24$, $SD = 2.64$; men: $M = 19.74$, $SD = 2.64$). Participants were recruited between 2016 and 2019 from the University of X's first year research participation scheme and were offered one credit for their participation in a study titled 'Speed-meeting Study'. To participate in the study, participants were requested to be 1)

heterosexual, 2) a native English speaker, 3) open to answering personal questions regarding their sexual history (for questions not relating to the current study), and 4) not in a committed relationship (required in 2017-2019). Participants who were known to each other (3.80%) or in a committed relationship (7.30%) were included in the main analyses; however, results with these participants excluded can be found in supplementary material E. Participants said yes to going on another date with their partner 46.54% of the time and they mutually said yes 20.95% of the time.

Before beginning, all participants were asked to read an information sheet which briefly detailed the procedure and highlighted the potential sensitivity of the sexually oriented questions. Participants were assured of confidentiality as well as being told at regular intervals that they may discontinue/omit answers without forgoing credit. They were then given an educational debriefing, including a debrief sheet. This study was approved by the Human Ethics Committee at the University of X (Ethics #16-PSYCH-4-65-JS).

Materials

Participants completed a series of questionnaires that were collected as part of a larger study investigating attraction. Only items included in the present study are detailed below.

Demographics. A range of demographic questions including age, sex, sexual orientation, and relationship status.

Speed-date ratings. Participants completed a 24-item questionnaire regarding each partner with whom they had a speed-date interaction. The first series of questions concerned the partner's personality attributes. Participants were asked to 'Please rate this partner on the following statements below' and were then presented with a statement regarding each trait individually, such as, for example, 'They are funny'. To ensure participants paid attention to the intelligence trait in particular, it was separated from the other traits and asked in the longer format of 'Thinking about this interaction, approximately how intelligent do you think this partner is?' The second series of questions concerned the partner's facial, bodily, and overall attractiveness (e.g. 'I would rate their overall attractiveness as…'). All questions in this section were rated on a 7-point response scale ranging from 1 = *Well Below Average* to 7 = *Well Above Average* with a midpoint of 4 = *Average*.

Verbal intelligence. To measure verbal intelligence, the latter (more difficult) half of Shipley's Vocabulary Scale was used (Zachary & Shipley, 1986). This scale included 20 items whereby the participant is presented with a target word (e.g. 'Jocose') and a series of four words (e.g. 'Humorous, Paltry, Fervid, Plain'). Participants are instructed that for each target word, they should 'please select the word that best matches its meaning'. These items

progressively become more difficult, beginning with well-known words such as 'Caption' and ending with more obscure words such as 'Temerity'.

Procedure

Pre-date. Four speed-date stations were constructed in the laboratory. Participants were seated opposite each other with Apple iPads so they were unable to see their partner's screen. Each station was separated by 1.7m room dividers to ensure the other couples were also unable to see their device screens. Upon arrival, participants were seated and given a participant information sheet. They were instructed to begin the pre-questionnaire if they agreed to participate. The pre-questionnaire consisted of demographics and other measures not used in this study. At the end of the pre-questionnaire, participants received on-screen instructions to wait quietly until others were finished.

Speed-dating. Once all participants had completed the pre-questionnaire, they were verbally instructed that they would now be given three minutes to interact with an opposite sex partner. Participants were instructed to speak about any topic until they heard a bell which would indicate the date had ended. After hearing the bell, participants were then instructed to begin completing the survey regarding their partner (as outlined in the speed-date ratings section of Materials). All participants were reminded to hold the iPads up to avoid their partner seeing the screen. Experimenters supervised the room to determine when all participants had finished completing ratings. At that point, the rotating sex (counterbalanced) moved onto the next station to start their next date. The process was then repeated until all opposite-sex dyads had interacted. If there was an uneven ratio of men and women, the extra participant(s) were instructed to sit quietly for three minutes during that round. In total, there were 123 speed-dating sessions with 729 participants. Participants participated in 2-5 dates ($M = 3.01$).

Post-date. Once all speed-dates and ratings had been completed, participants began completing the post-questionnaire which consisted of Shipley's Vocabulary Scale (Zachary & Shipley, 1986). Participants completed the first two sections and were instructed to wait quietly until all others had finished.

Statistical analysis

The nature of the design (i.e. participants rating multiple partners) creates dependencies in the data. The rating from each interaction between two people (Level 1) is cross-classified within both the participant receiving the rating (Level 2), and the partner who gave the rating (Level 2), all of which is nested within the session they both attended (Level 3). Therefore, it is necessary to use multilevel modeling (MLM) to account for the hierarchical structure of the

data. MLM analyses with partner ratings of attractiveness and intelligence at Level-1 and measured intelligence at Level-2 were used to evaluate main effects. Additionally, random slopes were included for all main effect variables (e.g. measured intelligence) for the relevant grouping factors (i.e. participant, and/or partner) to allow the slope between the independent and dependent variable to vary by group; however, these random slopes were removed when necessary to resolve convergence issues.

Results

Study 1

Target's intelligence level. Using eight intelligence subtests, we assessed our targets' measured intelligence (see Table S2). Results of cognitive ability tests are substantially intercorrelated, yielding a latent, general factor of intelligence, referred to as the *g factor* (Plomin & Deary, 2015). We conducted a principal component analysis and found that the first unrotated factor, the *g* factor, explained 37% of variance. This factor served as the criterion measure of the target's measured intelligence adopted in study 1.

Accuracy of intelligence perception. To investigate the accuracy of intelligence perception, we first correlated targets' *g* factor with an aggregated value of perceived intelligence using a Pearson product-moment correlation, $r = .34$, ($p < .001$; 95% CI [.14; .51], Fig. 2A). Aggregated perceiver values are commonly used in accuracy research; however, aggregates tend to lead to inflated accuracy estimates (Back & Nestler, 2016) and should be interpreted with caution. Therefore, we also used disaggregated ratings to determine the accuracy of individual women's judgments of intelligence ($\beta = .22$, $p < .001$, 95% CI [.07; .28]) in a structural equation model with standard errors clustered by target, modelling *g* as a hierarchical latent variable to correct for measurement error (see S3A). The results from both methods support our first prediction, suggesting that women are able to perceive intelligence with some degree of accuracy based on our three cues (cue 4: videos of men reading newspaper headlines aloud, cue 5: performing a pantomime task and cue 6: trying to make the experimenter laugh).

Fig. 2.
Aggregated perceiver accuracy for intelligence as measured by the g factor.
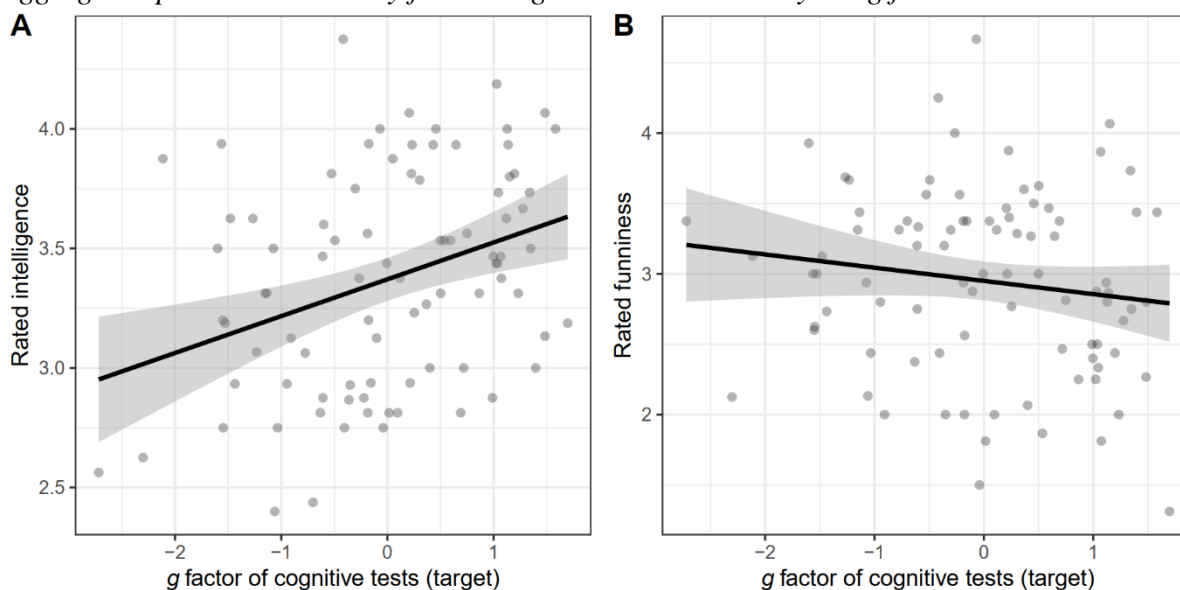
Note. The shaded area in grey reflects the 95% HDI.

Ratings of mate appeal. Women rated men's mate appeal operationalised as men's attractiveness as a short-term mate and long-term mate; however, we found that these ratings were highly correlated ($r = .92$). Therefore, all results are reported based on short-term mate attractiveness (henceforth referred to as sexual mate appeal); results for long-term mate attractiveness can be found in our supplement (see S3B).

Preference for funniness and perceived intelligence. If funniness is a display of intelligence, we would expect a relationship between men's measured intelligence and women's perception of men's funniness. Women's perception of men's funniness was associated with their perception of men's intelligence ($b = .30$, $p > .001$, 95% CI [.24; .36]). But contrary to expectations, measured intelligence was not associated with perceived funniness ($r = -.14$, $p = .18$, 95% CI [-.34; .07], Fig. 2B).

Further, we investigated whether funniness influences men's sexual mate appeal incremental to measured intelligence (Table 1). More intelligent men were rated to have a slightly lower sexual mate appeal ($b = -.14$, $p = .03$, 95% CI [-.26; -.01]), contrary to expectations. However, men who were perceived to be funnier had a higher sexual mate appeal ($b = .35$, $p < .001$, 95% CI [.26; .45]). These results do not support the notion that funniness is a display of intelligence. We found that men who were perceived to be more intelligent also had a higher sexual mate appeal ($b = .17$, p = .002, 95% CI [.06, .29]) (Table S11).

Table 1.

LM coefficients for associations between measured intelligence, humour and sexual mate appeal

	sexual mate appeal		
Term	Estimate	*p*	95 % CI
Intercept	.69	<.001	[.35; 1.03]
g factor	-.14	.03	[-.26; -.01]
Funniness	.35	<.001	[.26; .45]
Physical attractiveness	.24	<.001	[.17; .30]

Note. 88 Targets were rated by *n* = 30 women rating men's sexual mate appeal, *n* = 16 women rating men's funniness and *n* = 19 women rating men's physical attractiveness. The association of sexual mate appeal and g factor is depicted in Table S10.

Preference for more intelligent men. Contrary to our prediction that women would prefer more intelligent men, we found that more intelligent men were rated to have a slightly lower sexual mate appeal (*g* factor: *b* = -.07, 95% HDI [- .11, -.03]). Men's physical attractiveness was the main predictor of sexual mate appeal (*b* = 1.15, 95% HDI [1.05; 1.24]) (see Table 2). These findings do not support our second prediction, suggesting that women did not find intelligent men more appealing.

Adding initial intelligence cues. We predicted that more intelligent men's sexual mate appeal would increase more than it would for less intelligent men when shifting from only physical attractiveness information being available (cue 1-3; various physical and vocal attractiveness cues) to provision of additional cues related to men's intelligence (cue 4; reading newspaper headlines, which Borkenau et al. (2004) have found to be a task strongly related to accurate intelligence perception). As can be seen in Fig. 3, after cue 4 was presented, the increase in men's sexual mate appeal ratings did not depend on their intelligence (*g* factor x cue 4: *b* = .01, 95% HDI [-.02; .04]). This finding does not support our prediction, in that cues of intelligence did not uniquely contribute to sexual mate appeal ratings. Additionally, we predicted that further adding information on men's funniness (cue 6; make experimenter laugh) would provide a greater increase in sexual mate appeal for more intelligent men. Cue 5 (pantomime) was not presented in order to reduce test fatigue.

Contrary to our predictions, we found that the increase in men's sexual mate appeal did not depend on their intelligence (*g* factor x cue 6: *b* = .02, 95% HDI [-.02; .04]). Taken together with the previous finding, this casts further doubt on the notion that intelligence is attractive in men.

Fig. 3.
The aggregated sexual mate appeal ratings made after seeing each cue (or set of cues) was adjusted for physical attractiveness.
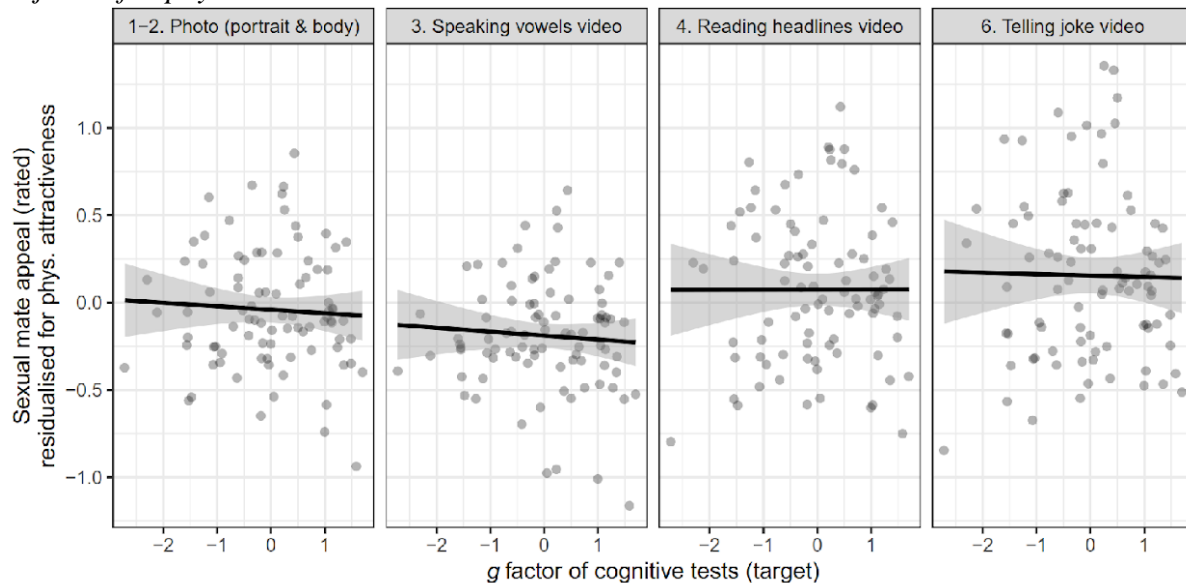
Note. The points shown in this plot show sexual mate appeal residualised for physical attractiveness. The shaded area in grey reflect the 95% HDI. The plot shows the slope of a linear regression predicting sexual mate appeal from the measured *g* factor. Intelligent men were not rated more favourably, even after intelligence-relevant information became available.

Table 2

Associations between sexual mate appeal and measured intelligence in sequential cue presentation

Term	sexual mate appeal	
	Estimate	95 % HDI
Intercept	.44	[0.09; 0.80]
Cue 1&2	-.30	[-.39; -.21]
Cue 4	.16	[.09; .23]
Cue 6	.36	[.25; .48]
physical attractiveness	1.15	[1.07; 1.22]
g factor	-.07	[-.10; -.03]
Cue 1&2 * physical attractiveness	-.10	[-.13; -.07]
Cue 4 * physical attractiveness	.07	[.04; .10]
Cue 6 * physical attractiveness	.14	[.11; .18]
Cue 1 & 2 * *g* Factor	-.01	[-.03; .02]
Cue 4 * *g* Factor	.01	[-.02; .04]
Cue 6 * *g* Factor	.02	[-.01; .05]

Note. Estimates and highest density intervals (HDI) from a Bayesian mixed effects location-scale model. Here, we show only the relevant non-varying effects on the mean, see Appendix S3B/online supportive materials for further control variables, varying effects and effects on scale. The reference category of the cue variable was set to the 'Vowels' video (cue 3), so that the interaction between cue 4 and measured intelligence captures the change in association at the point at which intelligence becomes task-relevant.

Additionally presented cues and attractiveness. Though the previous two results showed that change in sexual mate appeal with additional cues did not depend on men's intelligence, it should be noted that men's rated sexual mate appeal increased after cue 4 was presented (cue 4: b = .16, 95% HDI [.07; .24]) and further after cue 6 was presented (cue 6: b = .36, 95% HDI [.23; .50]). This raises the question of what other factor(s) involved in sexual mate appeal judgments were revealed in these later cues. We found that the increase in sexual mate appeal with additional stimuli was greater for more physically attractive men, with their ratings improving when after the presentation of cue 4 (cue 4 x physical attractiveness: b =

.07, 95% HDI [.04; .11]) and the presentation of cue 6 (cue 6 x physical attractiveness: $b =$ .14, 95% HDI [.10; .19]). Therefore, more physically attractive men did not only have a higher mate appeal, but they also benefited more from the later cues than did less physically attractive men.

Study 2

As predicted, more intelligent people were perceived to be more intelligent by their interaction partner, suggesting that intelligence is detectable in short live interactions ($\gamma =$ 0.08, 95% CI [.03; .13], $p = .002$). After aggregating ratings across raters, the correlation was $r = .12$ (Fig. 4). However, contrary to predictions, more intelligent people were not more likely to be rated as funnier by their partners ($\gamma = -0.01$, 95% CI [-.06; .05], $p = .841$). We found no evidence that the associations between intelligence and perceptions differed by sex (ps > .91).

As predicted, men perceived to be more intelligent or funnier were also rated as having a higher mate appeal by their interaction partners. However, measured intelligence did not predict rated mate appeal (Table 3, Fig. 5). We found no evidence that the associations with mate appeal differed by sex (ps > .18). Full results including random effects and moderation by sex can be found in the supplementary material F. Additionally, this pattern of results remained when controlling for both facial and bodily attractiveness, though some relationships between rated variables were attenuated. These results can be found in the supplementary material G.

Fig. 4.
Association between intelligence, as measured by the Shipley Institute of Living Scale (Vocabulary Subscale), and rated intelligence, after aggregating across raters.
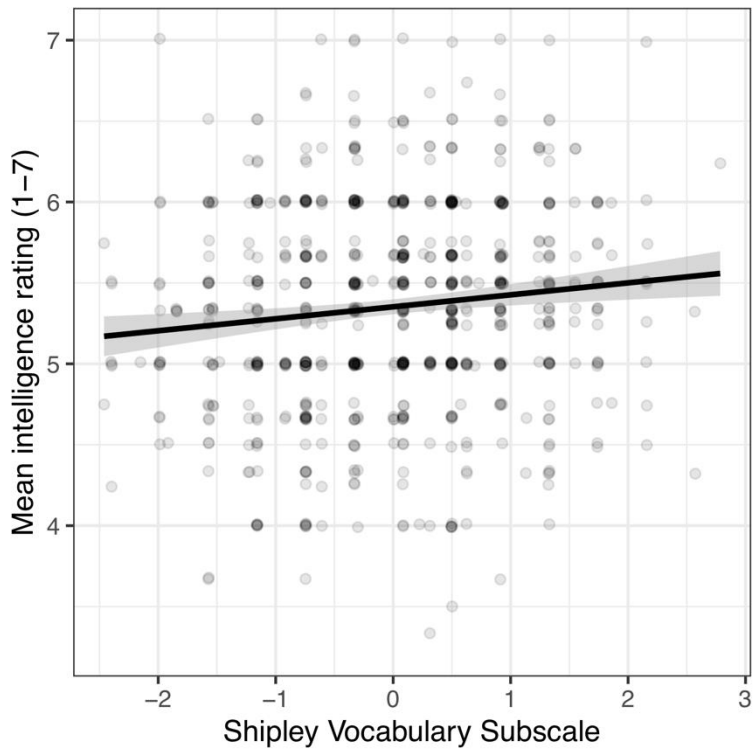
Note. Varying opacity of the dots is caused by overlap of multiple participants.

Fig. 5.
The association between intelligence, as measured by the Shipley Institute of Living Scale (Vocabulary Subscale), and rated mate appeal, after aggregating across raters.
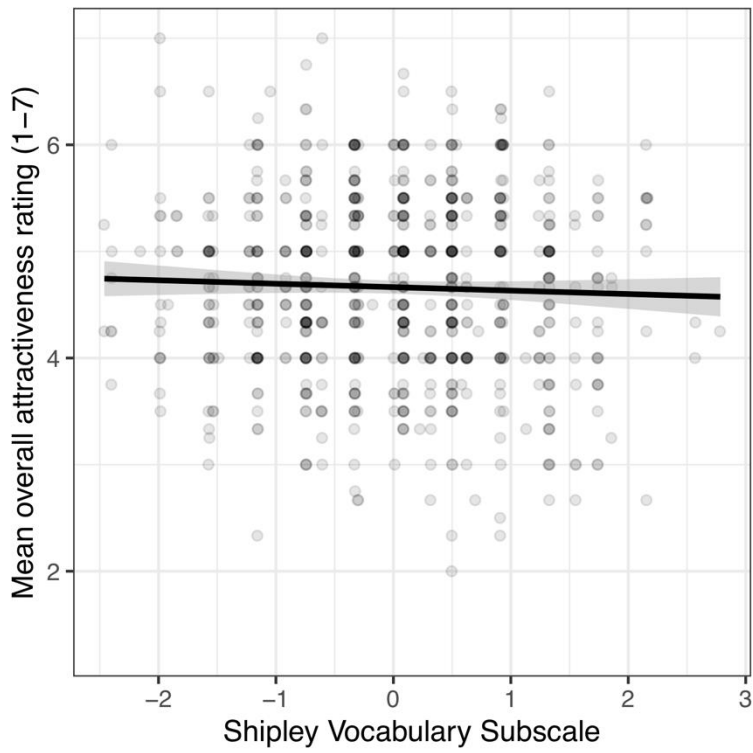
Fig. 6.
The association between intelligence, as measured by the Shipley Institute of Living Scale (Vocabulary Subscale), and funniness, after aggregating across raters.
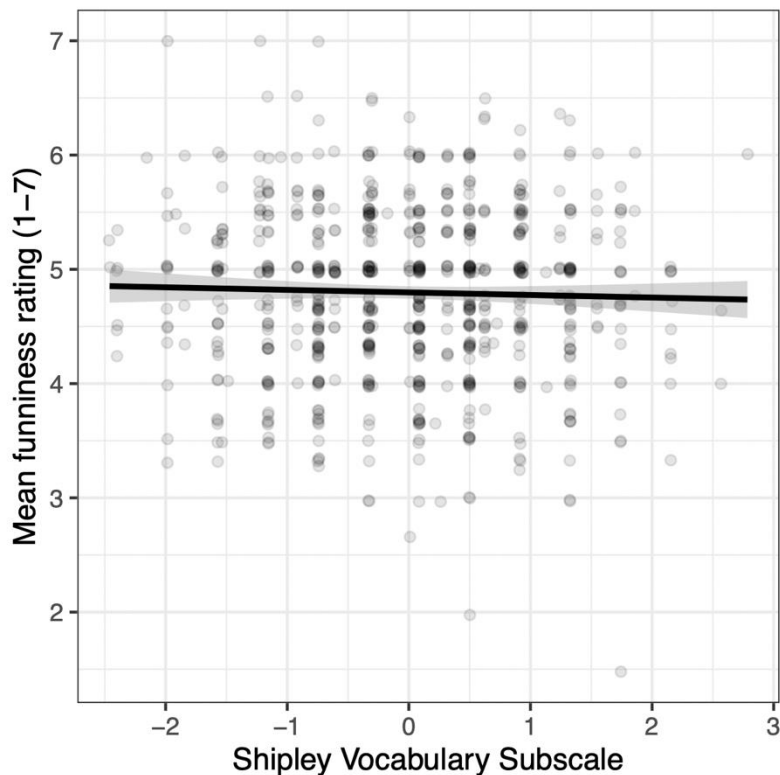
Table 3.
MLM coefficients for associations between the Shipley Institute of Living Scale (Vocabulary Subscale), rated intelligence, rated funniness, and rated mate appeal.

	Rated Mate Appeal (1-7)				
Predictors	*Estimates*	*CI*	*p*	$N_{interactions}$	$N_{participants}$
Shipley (Vocabulary Subscale)	-0.02	-0.07 – 0.04	.560	2114	727
Rated Intelligence	0.29	0.25 – 0.34	<.001	2118	728
Rated Funniness	0.41	0.37 – 0.44	<.001	2118	728

Note. Separate models were used for each predictor. In all models, sex was controlled. Full models are included in supplementary material E.

Discussion

The sexual selection theory of human intelligence proposes that intelligence evolved at least partly as a fitness indicator. Under this scenario, we would expect intelligence to be

sexually attractive to members of the opposite sex (Miller, 2000b, 2000a). Although intelligence is considered a highly attractive trait in a hypothetical partner (Buss et al., 1990; Li et al., 2002), it is less clear whether objectively assessed intelligence is indeed found attractive when evaluating a prospective partner. Studies directly assessing the link between intelligence scores and mating success are scarce and have inconsistent findings: Greengross and Miller (2011) found a positive association of women's ($r = .23$) but not men's ($r = .05$) verbal intelligence with a factor representing mate quantity, while in the UK Biobank (Neale Lab, n.d.) there is a negative genetic correlation of men's fluid intelligence and their number of sexual partners ($r = -.18$, $p < .001$) but no significant genetic correlation for women ($r = .07$, $p = .06$). In any case, these mate quantity measures ignore mate quality; our test in this paper of whether intelligence is found attractive is perhaps the more direct test of the sexual selection theory of human intelligence.

Our results replicate past findings (Borkenau et al., 2004) in showing that intelligence can be judged with above chance accuracy by members of the opposite sex at zero acquaintance. In the more ecologically valid setting of study 2, the association between actual and perceived intelligence is still significant, though attenuated. Taken together, these findings indicate that intelligence can be judged with above chance accuracy by members of the opposite sex at zero acquaintance.

Contrary to our hypotheses, more intelligent people were not rated as more appealing mates. Instead, only perceived intelligence was associated with higher mate appeal ratings. This finding illustrates the importance of using *measured* intelligence. Because rated physical attractiveness and perceived intelligence were strong predictors of mate appeal while measured intelligence was not, a halo effect could play a role. It is well established that physically attractive individuals are perceived as better in other socially desirable domains, independently of objective differences (Langlois et al., 2000). By gradually increasing the intelligence information and estimating the effect of intelligence beyond what can be observed from only physical cues, we could isolate the effect of information about intelligence, without the halo effect of physical attractiveness or any effects that intelligence might have on cues such as clothing or body shape. Contrary to our hypotheses, the increase in mate appeal after adding intelligence-related cues to visual and vocal attractiveness cues was not enhanced for more intelligent men.

One possibility is that invalid cues of intelligence are found attractive. Previous research has coded the frequency of different cues and their relationship with measured and perceived intelligence; a cue that is related to perceived intelligence and unrelated to

measured intelligence is necessarily an invalid cue. Reynolds and Gifford (2001) adopted this technique and showed that speech fluency was associated with greater perceived but not with measured intelligence. As people can detect intelligence to some extent, valid cues of intelligence are clearly perceptible. This is supported by studies finding cues that are associated with both measured intelligence and perceived intelligence (Murphy et al., 2003; Reynolds & Gifford, 2001). Had we evolved to find intelligence attractive because it signals genetic quality, we would have evolved to find valid cues of intelligence attractive. This pattern of results is not consistent with Miller's (2000a) proposal that intelligence acts as a fitness indicator.

Another possibility is that intelligence and related constructs are associated with positive outcomes across all environments. Therefore, people in these environments (i.e. cultures) will learn to associate intelligence with positive outcomes and, as a consequence, will report intelligence as being desirable. Previous research has shown that people believe intelligent individuals possess socially desirable traits such as being more competent and open-minded (Murphy et al., 2001). Choosing a competent mate in particular entails direct (i.e. non-genetic) fitness benefits related to resource provisioning, including income, socioeconomic status, and health, all of which are robustly predicted by intelligence (Deary, 2012). Since intelligence is also highly heritable, choosing a mate based on intelligence will also, as an indirect (genetic) benefit, pass on intelligence to the offspring. However, Miller (2000a, 2000b) goes further and predicts that intelligence evolved as a genetic fitness indicator that is preferred during mate choice for its indirect benefits (i.e., good genes sexual selection). If that was the case, intelligence should be sexually attractive, as partners who are found attractive for purely sexual encounters can only provide indirect, but not direct benefits. Of course, partners for exclusively sexual encounters can be chosen both during initial encounters with unacquainted strangers and from one's well-acquainted social surroundings (as was probably the more common case in our evolutionary past). But since intelligence is already accurately perceivable during initial encounters, as we and others have shown, it should already be found sexually attractive during such initial encounters if it had evolved as a fitness indicator through good genes sexual selection. Our finding that intelligence is not appealing during initial encounters despite being accurately perceivable suggests that intelligence is not a sexually attractive indicator of genetic quality, but rather preferred during later stages of long-term relationship formation (see Miller & Todd, 1998), probably due to its accompanying direct benefits.

According to Miller (2000b, 2000a), our ancestors would have used interpersonal humour during courtship to advertise and evaluate underlying intelligence and ultimately genetic quality. We found that ratings of funniness were associated with ratings of mate appeal, but contrary to our hypotheses and previous work (Greengross & Miller, 2011; Howrigan & MacDonald, 2008), measured intelligence did not predict ratings of funniness. Our measures of humour relied on being funny during a live interaction which presumably tapped into interpersonal humour, with its real-time evaluation and non-verbal cues. The more abstract tasks in previous research (Greengross & Miller, 2011; Howrigan & MacDonald, 2008) may have tapped skills that are more related to intelligence (e.g. drawing and writing) but not important for interpersonal humour.

In terms of limitations, study 1 and 2 used complementary approaches, with drawbacks of one study being addressed by strengths of the other study. Study 1 prioritised precision in our estimates of intelligence and a high degree of control over intelligence information at the expense of ecological validity, whereas study 2 did the opposite. A major limitation of study 1 was that ratings of men's sexual mate appeal were generally low, so that it seems unlikely that many of the men in our sample would have been chosen as partners by our raters. But in study 2, ratings of mate appeal were higher and many participants indicated hypothetical interest in going on a real date with their partner (for women 43.6% and for men 47.5% of interactions). Another limitation of study 1 was that women only saw short video sequences. At this initial stage of courtship, physical attractiveness is the most influential. This issue is partly addressed in study 2 in which participant's interactions are more reflective of a real courtship situation; however, we are still limited to the initial phase of getting acquainted. Still, the fact that participants could detect intelligence but were not influenced by it in their ratings of mate appeal calls into question the idea that intelligence is a fitness indicator.

A limitation of study 2 was that ratings of intelligence could be contaminated by cues about income (e.g. clothing and accessories). This limitation is mitigated in study 1 by showing images and voice prior to video content and controlling for these previous ratings when testing for the association between intelligence and mate appeal. Study 2 is limited by a less precise measurement of intelligence, but in study 1 we calculated a *g* factor based on multiple intelligence tests, thereby greatly increasing the reliability and validity of the intelligence construct. Relatedly, intelligence scores in study 2 were based on a university sample that is more educated and likely has a higher socioeconomic status than the general population. We partly addressed this in study 1, which was based on individuals from

university and the broader population to provide more diverse backgrounds and likely more diverse intelligence scores (see Table S4). However, all targets in study 1 were literate and not intellectually disabled, which means that intelligence variation was still limited to some degree. It is possible intelligence is important in a mate only to the extent that it is not very low (Zebrowitz & Rhodes, 2004).

In conclusion, our results do not support Miller's proposal that human surplus intelligence was shaped by intersexual selection. If our intelligence was shaped by the romantic and sexual choices across generations, this legacy should not only be reflected in our stated preferences, but also in mate choices. Instead we found that measured intelligence did not influence mate appeal, neither directly nor indirectly through funniness. Given the caveats to our findings, future research should extend our work by sampling a broader variation of the spectrum of intelligence and following courtship over a longer term beyond the initial contact.

References

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin*, *111*(2), 256–274. https://doi.org/10.1037/0033-2909.111.2.256

Arslan, R. C., Tata, C. S., & Walther, M. P. (2019). formr: A study framework allowing for automated feedback generation and complex longitudinal experience sampling studies using R (Version version v0.17.21) [Computer software].

Back, M. D., & Nestler, S. (2016). Accuracy of judging personality. In Hall, Judith A. and Schmid Mast, Marianne and West, Tessa V. (Ed.), *The Social Psychology of Perceiving Others Accurately* (pp. 98–124). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781316181959.005

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Berger, S., Graham, N., & Zeileis, A. (2017). Various Versatile Variances [Computer software].

Borkenau, P., Mauer, N., Riemann, R., Spinath, F. M., & Angleitner, A. [Alois] (2004). Thin slices of behavior as cues of personality and intelligence. *Journal of Personality and Social Psychology*, *86*(4), 599–614. https://doi.org/10.1037/0022-3514.86.4.599

Bressler, E. R., & Balshine, S. (2006). The influence of humor on desirability. *Evolution and Human Behavior*, *27*(1), 29–39. https://doi.org/10.1016/j.evolhumbehav.2005.06.002

Bürkner, P. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, *80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Buss, D. M., Abbott, M., Angleitner, A. [A.], Asherian, A., Biaggio, A., Blanco-Villasenor A., . . . Yang, K.-S. (1990). International Preferences in Selecting Mates: A Study of 37 Cultures. *Journal of Cross-Cultural Psychology*, *21*(1), 5–47. https://doi.org/10.1177/0022022190211001

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., . . . Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, *76*(1), 1–32. https://doi.org/10.18637/jss.v076.i01

Condon, D. M., & Revelle, W. (2014). The International Cognitive Ability Resource: Development and initial validation of a public-domain measure. *Intelligence*, *43*, 52–64. https://doi.org/10.1016/j.intell.2014.01.004

Darwin, C. (1871). *The descent of man and selection in relation to sex*. London: John Murray.

Deary, I. J. (2012). Intelligence. *Annual Review of Psychology*. (61), 453–482. https://doi.org/10.1146/annurev-psych-120710-100353

Deary, I. J., Liewald, D., & Nissan, J. (2011). A free, easy-to-use, computer-based simple and four-choice reaction time programme: The Deary-Liewald reaction time task. *Behavior Research Methods*, *43*(1), 258–268. https://doi.org/10.3758/s13428-010-0024-1

Eastwick, P. W., Eagly, A. H., Finkel, E. J., & Johnson, S. E. (2011). Implicit and explicit preferences for physical attractiveness in a romantic partner: A double dissociation in predictive validity. *Journal of Personality and Social Psychology*, *101*(5), 993–1011. https://doi.org/10.1037/a0024061

Greengross, G., & Miller, G. F. (2011). Humor ability reveals intelligence, predicts mating success, and is higher in males. *Intelligence*, *39*(4), 188–192. https://doi.org/10.1016/j.intell.2011.03.006

Hawrylycz, M. J., Lein, E. S., Guillozet-Bongaarts, A. L., Shen, E. H., Ng, L., Miller, J. A., . . . Riley, Z. L. (2012). An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature*, *489*(7416), 391–399. https://doi.org/10.1038/nature11405

Howrigan, D. P., & MacDonald, K. B. (2008). Humor as a Mental Fitness Indicator. *Evolutionary Psychology*, *6*(4), 147470490800600. https://doi.org/10.1177/147470490800600411

Klasios, J. (2013). Cognitive traits as sexually selected fitness indicators. *Review of General Psychology*, *17*(4), 428–442. https://doi.org/10.1037/a0034391

Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin*, *126*(3), 390–423. https://doi.org/10.1037/0033-2909.126.3.390

Lehrl, S. (2005). MWT-B—Mehrfach-Wortschatz-Intelligenztest. *Göttingen, Germany: Hogrefe*.

Li, N. P., Bailey, J. M., Kenrick, D. T., & Linsenmeier, J. A. W. (2002). The necessities and luxuries of mate preferences: Testing the tradeoffs. *Journal of Personality and Social Psychology*, *82*(6), 947–955. https://doi.org/10.1037/0022-3514.82.6.947

Miller, G. F. (2000a). *The mating mind: How sexual choice shaped the evolution of human nature*. New York, NY: Doubleday.

Miller, G. F. (2000b). Mental traits as fitness indicators - Expanding evolutionary psychology's adaptationism. In D. LeCroy & P. Moller (Eds.), *Annals of the New York Academy of Sciences: Vol. 907. Evolutionary perspectives on human reproductive behavior* (pp. 62–74). New York: New York Academy of Sciences. https://doi.org/10.1111/j.1749-6632.2000.tb06616.x

Miller, G. F., & Todd, P. M. (1998). Mate choice turns cognitive. *Trends in Cognitive Sciences*, *2*(5), 190–198. https://doi.org/10.1016/S1364-6613(98)01169-3

Mink, J. W., Blumenschine, R. J., & Adams, D. B. (1981). Ratio of central nervous system to body metabolism in vertebrates: its constancy and functional basis. *The American Journal of Physiology*, *241*(3), R203 - R212. https://doi.org/10.1152/ajpregu.1981.241.3.R203

Murphy, N. A., Hall, J. A., & LeBeau, L. S. (2001). Who's smart? Beliefs about the expression of intelligence in social behavior. *Representative Research in Social Psychology*, *25*, 34–42.

Murphy, N. A., Hall, J. A., & Randal, C. C. (2003). Accurate intelligence assessments in social interactions: Mediators and gender effects. *Journal of Personality*, *71*(3), 465–493. https://doi.org/10.1111/1467-6494.7103008

Neale Lab UKBB (n.d.). *UKBB Genetic Correlation*. https://ukbb-rg.hail.is/rg_browser/

Peirce, J. W. (2007). Psychopy--Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1-2), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Plomin, R., & Deary, I. J. (2015). Genetics and intelligence differences: Five special findings. *Molecular Psychiatry*, *20*(1), 98–108. https://doi.org/10.1038/mp.2014.105

Prokosch, M. D., Coss, R. G., Scheib, J. E., & Blozis, S. A. (2009). Intelligence and mate choice: Intelligent men are always appealing. *Evolution and Human Behavior*, *30*(1), 11–20. https://doi.org/10.1016/j.evolhumbehav.2008.07.004

Puts, D. A. (2010). Beauty and the beast: Mechanisms of sexual selection in humans. *Evolution and Human Behavior*, *31*(3), 157–175. https://doi.org/10.1016/j.evolhumbehav.2010.02.005

R Core Team (2019). R: A language and environment for statistical (Version 3.6.0) [Computer software]. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Reynolds, D. J., & Gifford, R. (2001). The Sounds and Sights of Intelligence: A Lens Model Channel Analysis. *Personality & Social Psychology Bulletin*, *27*(2), 187–200. https://doi.org/10.1177/0146167201272005

Ronay, R., & Hippel, W. v. (2010). The Presence of an Attractive Woman Elevates Testosterone and Physical Risk Taking in Young Men. *Social Psychological and Personality Science*, *1*(1), 57–64. https://doi.org/10.1177/1948550609352807

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*(2), 1–36. Retrieved from http://www.jstatsoft.org/v48/i02/

Schipolowski, S., Wilhelm, O., Schroeders, U., Kovaleva, A., Kemper, C. J., & Rammstedt, B. (2013). BEFKI GC-K: eine Kurzskale zur Messung kristalliner Intelligenz. *Methoden, Daten, Analysen (Mda)*, *7*(2), 153–181. https://doi.org/10.12758/mda.2013.010

Sprecher, S., & Regan, P. C. (2002). Liking Some Things (in Some People) more than Others: Partner Preferences in Romantic Relationships and Friendships. *Journal of Social and Personal Relationships*, *19*(4), 463–481. https://doi.org/10.1177/0265407502019004048

Treffenstaedt, C., & Wiemann, P. (2018). Alfred - a library for rapid experiment development (Version 0.2b5) [Computer software].

Wilbur, C. J., & Campbell, L. (2011). Humor in romantic contexts: Do men participate and women evaluate? *Personality & Social Psychology Bulletin*, *37*(7), 918–929. https://doi.org/10.1177/0146167211405343

Zachary, R. A., & Shipley, W. C. (1986). *Shipley institute of living scale: Revised manual*. WPS, Western Psychological Services.

Zebrowitz, L. A., & Rhodes, G. (2004). Sensitivity to "bad genes" and the anomalous face overgeneralization effect: Cue validity, cue utilization, and accuracy in judging intelligence and health. *Journal of Nonverbal Behavior*, *3*(28), 167–185. https://doi.org/10.1023/B:JONB.0000039648.30935.1b

Zeileis, A. (2004). Econometric Computing with HC and HAC Covariance Matrix
Estimators. *Journal of Statistical Software*, *11*(10), 1–7.
https://doi.org/10.18637/jss.v011.i10

Zeileis, A., & Hothorn, T. (2002). Diagnostic Checking in Regression Relationships. *R News*,
2(3), 7–10.

Author Contributions

JMH, LP, RCA planned and supervised study 1. JCD and MD conducted parts of study 1. RCA and JCD analysed study 1 with consultation from PCB. BPZ and MJS planned study 2. MJS partially conducted, supervised student data collection for, and analysed study 2. MJS, JCD, BPZ and RCA wrote the manuscript. All authors read and approved the final manuscript.

Competing Interests statement

All authors declare no competing interests.

Open Practices Statement

Study 1 was preregistered on the Open Science Framework https://osf.io/qt9cw/) All analysis scripts have been made publicly available at the Open Science Framework (https://osf.io/rs3tg)

Funding

Study 2 was supported by the Australian Research Council [FT160100298].

Supplement (Manuscript 1)

S1. Table with deviations from pre-registration

Study 1 was pre-registered as part of three different theses. All theses had a slightly different focus. After the preregistration, the project continued to develop and we collected additional rating data to address methodological shortcomings. Throughout the process, we realised that some preregistered hypotheses were insufficient. We also opted for more appropriate analyses. In the following all deviations from the pre-registration are outlined.

Deviations in our hypotheses

Preregistration	Manuscript	Explanation/Solution
Thesis A focuses on the association between intelligence and attractiveness and made the following predictions: A1. More intelligent men are preferred as long-term mates by women. A2. Men who are perceived as more intelligent are preferred as long-term mates by women. **A3. More intelligent men are preferred as short-term mates by women.** **A4. Men who are perceived as more intelligent are preferred as short-term mates by women.** A5. Perceived creativity is predicted by intelligence and perceived intelligence. A6. Men who are perceived as more creative are preferred as long-term mates, incremental to intelligence.	We derived the following hypotheses in our manuscript: 1. Women's intelligence ratings for male targets will be positively correlated with targets' psychometrically measured intelligence. 2. Men's measured intelligence will be significantly positively correlated with women's attractiveness ratings. 3. Perceived funniness is associated with measured intelligence. Men's funniness and men's perceived intelligence predicts their rated mate appeal above and beyond the effect of their intelligence.	All preregistered hypotheses in bold are still part of our manuscript (although phrased differently). Differences are: 1. All hypothesis on long-term mate ratings are not part of our main manuscript anymore, however reported in our supplement. This deviation results from the fact that short-term mate and long-term mate ratings were highly correlated ($r = .92$). In our design participants do not seem to differentiate much between short-term mate and long-term mate ratings. Hence, we cannot test for differences between the two outcomes. 2. None of the hypotheses of thesis C are included in the current manuscript. Thesis C focused on the accuracy of personality judgments and do not fit the scope of this paper.

A7. Men who are perceived as more creative are preferred as short-term mates, incremental to intelligence. **A8. Perceived sense of humor is predicted by intelligence and perceived intelligence.** A9. Men who are perceived as more humorous are preferred as long-term mates, incremental to intelligence. **A10. Men who are perceived as more humorous are preferred as short-term mates, incremental to intelligence.** **A11. Women can accurately assess men's intelligence based on thin slices of behavior.** A12. Narcissism and shyness may moderate the effect of intelligence on men's appeal to women as long- and short-term mates, as they are expected to have an effect on intelligence displays and their perception at zero acquaintance independent of actual target intelligence. Similarly, thesis B focuses on the association of intelligence and attractiveness. Though, the thesis goes beyond the former in a more detailed investigation of the relationship between the two. B1. There will be a significant change in the rating of men´s attractiveness as short-term and long-term mates after shifting from mere visual and vocal attractiveness information (full body photo, video of reading vowels aloud) to cues also indicative	4. The increase in men's short-term mate attractiveness after adding additional cues related to intelligence (i.e. reading newspaper headlines; making experimenter laugh) will depend on men's intelligence, such that the attractiveness increase will be greater for more intelligent men.	3. Similarly, we did not include any hypothesis on Narcissism, shyness and creativity for the sake of brevity and since we found no main effect of intelligence on attraction.

of intelligence (reading headlines aloud, telling experimenter something funny). Effects are expected to be more pronounced when it comes to long-term mating, but if they are also found for short-term mating this can be interpreted as evidence for intelligence as a cue to genetic quality.

B2. The more cues indicative of intelligence are added, the larger the change in appeal ratings is expected to be. Appeal ratings should increase more for more intelligent men, and more so after the fourth (telling something funny) than after the third rating (headlines). These intelligence-dependent increases should occur for both long-term and short-term mate ratings, but more so when it comes to long-term mating.

B3. More intelligent men will be rated as more desirable short-term and long-term mates.

Thesis C focused on the accuracy of intelligence and personality judgements.

C1. There will be a positive correlation between self-reports and observer ratings of the Big Five dimensions. Correlations are expected to range between r = 0.2 to 0.4.

C2. The correlations between self-reports and observer ratings will be higher for extraversion and conscientiousness than for neuroticism, openness to experience, and agreeableness

C3. Measured psychometric intelligence and self-rated openness to experience are strong predictors for observer-rated creativity and observer-rated humour production ability. C4. Attempting to replicate the results of a study by Murphy and colleagues (2003), the ratings of intelligence made by female raters are predicted to be more accurate (more highly correlated with measured psychometric intelligence) than the ones made by male raters.		

Deviations in our recruitment

Preregistration	Manuscript	Explanation/Solution
Recruitment of targets was limited: ● to 80 male participants ● men who report a heterosexual orientation ● who are aged between 18 - 30	Recruitment slightly differed: ● 88 target men were recruited ● 2 target men reported a bisexual orientation and 1 target reported to be homosexual ● age ranged from 18 to 31 years	In a robustness check, we excluded targets reporting a non-heterosexual orientation targets. Results are reported in our appendix S3.
Recruitment of 55 female raters: ● who report a heterosexual orientation ● 10 women rate men's physical attractiveness (Rating Study 1.1) as well as men's humour and intelligence (Rating Study 1.2)	179 female rater were recruited: ● 1 reported a bisexual orientation ● 19 women participated in Rating Study 1.1 (physical attractiveness) ● a different set of 16 women participated in Rating Study 1.2 (rating of intelligence and humour)	A total of 179 female raters were recruited. At the time point when writing the preregistration, only Rating Study 1.1 - 1.4 were intended to take place. We decided to run Rating Study 1.5 and 1.6 post-hoc to address methodological shortcomings of study 1.4. Hence, only Rating Study 1.1 - 1.4 should be taken into account when comparing differences between the manuscript and the preregistration. Thus, the number of

• 10 women rate men's STM and LTM attractiveness (Rating Study 1.3) • 25 women rate men's STM and LTM after each presented cue (Rating Study 1.4)	• 30 women participated in Rating Study 1.3 (LTM and STM attractiveness) • 25 women participated in Rating Study 1.4 (STM/LTM after each Cue) • 30 women participated in Rating Study 1.5 (40 initially participated but 10 had to be excluded due to technical issues) • 59 women participated in Rating Study 1.6	preregistered to actually recruited female raters differs from 55 to 179. • It was preregistered that 10 women rate men's physical attractiveness, as well as men's intelligence and humour (Rating Study 1.1 and 1.2). To prevent potential halo effects, we recruited two different sets of female raters. One set rated men's physical attractiveness and the other set rated men's humour and intelligence (described in further detail under deviation in design) • The one woman reporting a bisexual orientation only rated men's intelligence and humour and made no attractiveness rating. Hence, we decided that we do not have to exclude her in any analysis.

Deviations in our design

Preregistration	Manuscript	Explanation/Solution
Among other measures, 24 items of the ICAR (International Cognitive Ability Resource; Condon & Revelle, 2014; German translation by our lab) will be used to measure the targets level of intelligence.	Instead of 24, a total of 25 items out of the ICAR (International Cognitive Ability Resource; Condon & Revelle, 2014; German translation by our lab) were adopted.	We included an additional item of the dimension verbal reasoning in order to increase reliability.
A full body and a facial photograph of each participant will be taken as a stimulus including cues on their physical appearance (face, body, posture, and style). These will be standardized for posture and neutral facial expression, but in normal street appearance.	Participants did not receive any instruction for posture or facial expression.	Targets did not receive an instruction to have a neutral facial expression and posture due to a miscommunication with our research assistants. However, we selected the most neutral picture out of our videos which led to semi-standardised pictures.

Female raters watch all three videos of a man and rate his physical attractiveness, intelligence and humour.	To rate men's physical attractiveness, female raters only saw a men's full body and facial picture. A different set of female raters rated men's intelligence and humour based on all three videos of them performing several tasks (video headlines, pantomime and make experimenter laugh).	To minimize potential halo-effects, one set of women rated men's physical attractiveness and a different set of women rated men's intelligence and humour. Physical attractiveness was rated based on the photographs and not on the videos since this is the standard procedure in the literature.

Deviations in our analyses

Preregistration	Manuscript	Explanation/Solution
To test whether women can accurately perceive men's intelligence (hypothesis 1), it was preregistered to correlate men's g factor with an aggregated value of perceived intelligence. Female ratings of each men are aggregated.	We correlated the aggregated value of perceived intelligence with men's actual intelligence. Additionally, we calculated the single perceiver accuracy. In a multi-level model, we predicted perceived intelligence ratings with men's actual intelligence. We specified a random effect for each men and each female rater.	Aggregated values inflate the accuracy because measurement error is reduced and because of a wisdom of crowds effect. Because people may mainly have only their own perception to go on, we also calculated the single perceiver accuracy (disattenuated for measurement error in the g factor, but not for measurement error in the rating).
To test whether more intelligent men are rated as more attractive as a STM mate (hypothesis 2), it was preregistered to regress men's actual intelligence onto their STM attractiveness while adding	Our main model (explained below under hypothesis 4) allows to answer the question whether more intelligent men are rated as more attractive as a potential partner. A	These analyses were only specified for Rating Study 1.3. Though, we also have the possibility to investigate the assumption in our analyses of Rating Study 1.4 - 1.6 without the need of

target's age, relationship status as well as experiment (dummy coded) as covariates onto the model. In a second model, instead of actual intelligence the aggregated perceived intelligence was added to the model.	positive main effect of *g* factor would be in line with our second hypothesis. In our supplement we included further analyses on hypothesis 2 based on ratings of rating study 1.3. In these analyses, we specified a linear model and predicted men's STM attractiveness with men's *g* factor, physical attractiveness. The difference is that we now used single ratings. However, we needed to correct standard errors as ratings were clustered in different sets of female raters. In a further model we repeated our analyses just like preregistered using aggregated ratings. We did not include target's relationship status, age and the experimenter as covariates into our model. Nevertheless, omitting these covariates did not change the results.	specifying a further model. To keep our main analyses lean, we only reported the main effect of men's *g* factor in our main model (based on sample 1.4 - 1.6). We shortly mention results of sample 1.3 in the main text but included a more detailed overview our appendix. Nevertheless, the preregistered model relies on aggregated ratings which inflate type I errors (DeBruine, 2019; Judd et al., 2016).
To test the influence of humour on attractiveness as a potential partner above the influence of intelligence (hypothesis 3), again it was preregistered to use aggregated ratings.	We included the preregistered model in our appendix. However, we specified a multilevel model in our manuscript.	We argue a multilevel model instead of aggregating ratings being the more appropriate way of analysis (as explained above).
To test whether the STM attractiveness increases for more intelligent men after presenting additional cues related to intelligence (hypothesis 4), a within-subject ANCOVA was preregistered. While measured intelligence should be added as a covariate to the model, target	We specified Bayesian models with weakly informative priors (for a detailed description see Method and S2C).	At the time of the preregistration, we did not intend to replicate our results in two further studies. However, our analyses have to take into account the varying influences of each rating study, explaining additional differences in our preregistered and actual analyses.

age, relationship status as well as experimenter should be added as control variables.		Additionally, we no longer consider an ANCOVA to be an appropriate way of analysing our data. A Bayesian model better satisfies the needs of our hierarchical data.

S2. Detailed description of Study 1

A. Materials

Intelligence measures. We adopted multiple measures to assess men's intelligence. In the online screening survey, we used the 16-item short version of the International Cognitive Ability Resource (ICAR, Condon & Revelle, 2014) which is a public-domain assessment tool to assess cognitive abilities. We enriched the short-version with 9 additional items of the long version to increase reliability. Hence, we assessed 4 dimensions namely verbal reasoning, matrix reasoning, letter and number series as well as mental rotation three-dimensional.

In the lab, we used three additional measures. We adopted the multiple choice vocabulary test [Mehrfachwahl-Wortschatz-Intelligenztest] (MWT, Lehrl, 2005) which is a measure to assess participants' general intelligence level, especially their crystallized intelligence.

Our third measure BEFKI GC-K [Kurzskala des Berliner Tests zur Erfassung fluider und kristalliner Intelligenz] (BEFKI GC-K, Schipolowski et al., 2013) also assesses participants' crystallized intelligence based on a 12-item knowledge scale.

Additionally, we adopt the Deary-Liewald reaction time task which is a computer-based reaction time programme (DLRT, Deary et al., 2011). The DLRT assesses simple reaction times (SRT) as well as four-choice reaction times (CRT). To assess the SRT, in each of the 20 runs participants pressed a button in response to a single stimulus. For the CRT, 4 stimuli were presented in 40 runs. In each run, participants had to press one button corresponding to the correct stimuli.

Table S2.

Mean values, standard deviations and ranges of intelligence measures

Variable	*M*	*SD*	min	max
BEFKI	9.52	1.81	4	12
CRT	450.90	47.15	351.2	557.1
SRT	289.8	21.38	244.60	337.20
ICAR letter	4.38	1.80	0	6
ICAR matrix	3.52	1.50	0	6
ICAR rotation	3.53	2.11	0	6
ICAR verbal	5.21	1.43	1	7
MWT	23.57	4.25	11	32

Note. Sum scores are reported for the MWT.

Table S3.

Comparison of men who dropped out during our online study with our final target sample.

	Targets	Dropout		Sample size (*n*)	
Variable	*M* (*SD*)	*M* (*SD*)	*p*-value	Target sample	Dropout sample
ICAR verbal	5.21 (1.43)	4.64 (1.43)	=.02	88	67
ICAR letter	4.38 (1.80)	3.69 (1.68)	=.03	88	48
ICAR matrix	3.52 (1.50)	3.00 (1.61)	=.09	88	38
ICAR rotation	3.53 (2.11)	2.89 (1.82)	=.09	88	36

Note. *M* = mean number of correct responses.

Table S4.

Range of IQ Scores.

	Male Targets			Normative Sample
	Targets *M* (*SD*)	[Min;Max]	average IQ [Min; Max]	*M* (*SD*)
MWT-B	24.90 (4.50)	[11.59;33.82]	95 [79; 130]	the manual provides an overview of the number of correct responses and each corresponding IQ score only
BEFKI	9.52 (1.81)	[4;12]	112 [80;126]	7.44 (2.61)
ICAR verbal	0.74 (0.20)	[0;1]	107 [83;116]	0.52 (.46)
ICAR letter	0.73 (0.30)	[0;1]	106 [84;114]	0.53 (0.49)
ICAR matrix	0.59 (0.26)	[0;1]	104 [86; 117]	0.46 (0.48)
ICAR rotation	0.59 (0.35)	[0;1]	112 [93;126]	0.22 (0.45)
SRT	289.81 (21.38)	[244.6;337.2]	-	243.1 (17.6)
CRT	450.86 (47.15)	[351.2;557.1]	-	388.0 (45.0)

Note. Note that we rescaled values of the MWT-B ranging from 1 to 35 on a scale ranging from 1 to 37. We report the average proportion of correct responses for the ICAR. All other values are the average number of correct responses. We calculated IQ-scores using the following formula: (M(target sample) - M(normative sample)/ SD(normative sample)) * 15 + 100.

B. Demographics

Study 1 comprises 6 rating studies which are described in more detail below. Participation was rewarded with course credit. Women could only participate in one of the six rating studies.

Rating study 1.1. In October and November 2014, we assessed the target's physical attractiveness. Hence, 19 women rated men's psychical attractiveness after seeing a facial and body photograph of our male targets (stimuli 1 and 2). Women were on average 23 years old (SD = 3.14, range = 18 - 28). 18 women reported to be students. All women reported a heterosexual orientation. 42.11% women were in a relationship.

Rating study 1.2. In October 2014, we also invited 30 participants who rated the targets' intelligence, funniness, creativity and personality after watching 3 videos of target men (cue 4 - 6). 16 of those raters were female (mean age = 21.06, SD = 3.44, range = 19 - 30) and 14 raters were male (mean age = 21.86, SD = 2.83, range = 19 - 29). For our purposes, only female ratings of men's intelligence and funniness are used in subsequent analysis. 15 women's highest level of education was a high school degree and 1 woman reported a university degree as her highest level of education. 87.5% of those women were in a relationship. 15 women reported a heterosexual orientation and 1 woman reported to be bisexual.

Rating study 1.3. In November 2014, we assessed participants' short-term and long term mate attractiveness. A new set of 30 women rated men's attractiveness as a short-term and long-term mate after watching 3 videos of target men (cue 4 - 6). All women were students and were on average 20.87 years old (SD = 2.42, range = 18 - 28). 36.67% of those women were in a relationship. All women reported to be heterosexual.

Rating study 1.4. From August to September 2014, 25 heterosexual women participated in this study (mean age = 23.96 years, SD = 2.82, range = 20 - 30). 60% reported having a high-school degree and 40% a university degree as their highest level of education. 60% of those women were currently involved in a romantic relationship.

Rating study 1.5. The first replication of study 1.4 took place between June 2015 and August 2015. We recruited through various online channels (e.g. Facebook, university platform), as well as posters on campus at the University of Goettingen. We aimed to recruit 30 raters. Participants had to be female and between 18 and 30 years old. Women had the possibility to either receive course credit or participate in a lottery as an incentive for taking part in our study. 40 women initially participated in our rating study. Though, as one woman was already familiar with the study, she was excluded from analysis. Due to technical problems nine further participants had to be excluded from analysis as they were only able to rate less than half of the targets leaving us with a final sample size of 30 women for final analysis (mean age = 21.80 years, SD = 2.70, range = 19 - 29) . From these datasets, 26 were completed while four datasets only contained at least two thirds of all ratings, again due to technical issues. 13% of those women stated to have a university degree and 87% reported having a high school degree as their highest level of education. 18 (60%) women were currently involved in a romantic relationship. One woman reported being bisexual, all other women reported being heterosexual.

Rating study 1.6. The recruitment of our second replication took place in January to February 2016 at the University of Leipzig. We recruited 59 female raters (mean age = 21.29

years, SD = 3.54, range = 18 – 36). 32 (54%) of those women were currently involved in a romantic relationship, with 88% having a high school degree and 12% a university degree as their highest level of education. 49 (83%) women reported a heterosexual orientation and 10 (17%) women a bisexual orientation. Nearly all women except four women were fluent in German.

Table S5.
Number of ratings, mean values, standard deviations and ranges in each rating study as well as Cronbach's α and interclass correlation coefficients (ICC) for single ratings.

Variable	single ratings							aggregated ratings			
	N	*M*	*SD*	Min	Max	*ICC*	*α*	*n*	*M*	*SD*	Min	Max
Rating Study 1.1											
Physical Attractiveness	1,657	2.82	1.53	1	7	.24	.93	88	2.82	0.79	1.42	5
Rating Study 1.2											
Perceived intelligence	1,368	3.37	0.87	1	5	.22	.91	88	3.37	0.45	2.40	4.38
Perceived funniness	1,368	2.95	1.07	1	5	.28	.93	88	2.95	0.65	1.31	4.67
Rating Study 1.3											
Short-term mate attractiveness	2,620	2.41	1.68	1	7	.36	.96	88	2.41	1.03	1.03	5.03
Long-term mate attractiveness	2,618	2.64	1.68	1	7	.27	.94	88	2.64	0.90	1.17	4.53
Rating Study 1.4 - 1.6											
Short-term mate attractiveness	33,358	-.01	2.25	-4.22	8.36	n/a	n/a				
Long-term mate attractiveness	33,358	-.01	2.18	-4.13	8.40	n/a	n/a				

Note. n/a = not applicable. *ICC2* = Intraclass correlation coefficients for a random set of judges who rate every target. α = Cronbach's alpha.

Table S6.

Correlation coefficients of target's g factor, age, aggregated ratings of their perceived creativity, humour and physical attractiveness.

Variable	1.	2.	3.	4.	5.
1. g factor	-				
2. creativity	-.02	-			
3. humour	-.14	.84***	-		
4. perceived intelligence	.34**	.50***	.40***	-	
5. physical attractiveness	-.26*	.44***	.59***	.11	-
6. age	.14	-.06	-.15	.21*	-.21*

Note. * $p < .05$, ** $p < .01$, *** $p < .001$

C. Further information on Bayesian models

Our Bayesian models were fitted in Stan (Carpenter et al. 2016) via brms (version 2.10.0, Bürkner, 2017). Because of the varying rating scales across studies and the possibility that raters would use scales differently, we used mixed effects location-scale models, which allowed not only the mean response but also the residual variance to differ by study, Cue, target, and rater. These more complex models fit better according to the approximative leave-one-out information criterion (LOO-IC), although the main results did not change. To improve convergence and sampling efficiency, we used weakly informative priors, specifically normal(M=0,SD=5) for the non-varying effects on the means, Cauchy(0,3) for the varying effects on the means, N(0,1) for non-varying and varying effects on the residual variation. We fit four parallel chains to assess convergence using the Rhat statistic. The full code of the models is documented in our repository on OSF.

S3. Robustness checks and additional analyses
A. Hypothesis 1: Single perceiver accuracy

To investigate whether women can accurately perceive men's intelligence, we correlated men's g factor with an aggregated value of their perceived intelligence. Though as aggregated values tend to inflate accuracy estimates, we also used disaggregated ratings to determine the accuracy of individual women's judgements of intelligence ($\beta = .18$, $p < .001$, 95% CI [.07; .28]) in a structural equation model, modelling g as a hierarchical latent variable to correct for measurement error (see Table S7). This model fit our data well χ^2 (24, 2581 observations clustered in 88 targets) = 421.47 , $p < .001$ (comparative fit index [CFI] = .919, normative fit index [NFI], standardized root mean square residual [SRMR] = .038, root mean squared error of approximation [RMSEA] = .08, 90% CI [.073, .087]). The results from both methods support our first predictions, suggesting that women are able to perceive intelligence with some degree of accuracy based on our three stimulus types.

Table S7.
Results of structure equation model on the accuracy of intelligence perception where the g factor is modelled as a hierarchical latent variable to correct for measurement error.

Term	Estimate	p	95 % CI
	Latent variables		
Reaction Time			
CRT	1.15	.40	[-1.51, 3.80]
SRT	0.34	.36	[-0.39, 1.06]
Language Tests			
BEFKI	0.50	.29	[-0.43, 1.43]
MWT	1.80	.31	[-1.62, 5.3]
ICAR			
Verbal	0.65	<.001	[0.33, 0.98]
Letter	1.04	<.001	[0.51, 1.56]
Rotation	1.01	<.001	[0.48, 1.55]
matrices	0.72	<.001	[0.32, 1.11]
g Factor			
ICAR	0.88	.05	[-0.01, 1.77]
Reaction Time	0.18	.50	[-0.34, .69]
Language	1.78	.41	[-2.47, 6.03]
	Regressions		
g ~ Intelligence response	0.22	.005	[0.07, .38]

B. Hypothesis 2 to 4: Results on LTM Attractiveness

As described, women's ratings of men's long-term mate and short-term mate attractiveness was highly correlated. Hence, results were extremely similar. We therefore decided to report only results on short-term mate attractiveness in our main manuscript and report results on long-term mate attractiveness as part of our supplement.

We assumed that more intelligent men were rated as more attractive as a potential partner. For short-term mate attractiveness, we found a contradicting effect (Table S14): more intelligent men were rated as less attractive as a short-term mate (b = -.07, 95% HDI [-.11; -.03]). For long-term mate attractiveness, we found no association between g and long-term mate attractiveness (b = -.02, 95% HDI [-.06; .01]).

With the previous results being based on female raters participating in study 1.4 - 1.6, we found similar results in a second set of raters. In study 1.3 women also rated target's short-term and long-term mate attractiveness. The difference is that women saw cue 4 (pantomime) instead of cue 3 (vowels) and rated men's mate appeal only once after watching all three videos. In this second set of female raters, we investigated our second hypothesis namely whether more intelligent men had a higher mate appeal once using aggregated (Table S8) and second using single ratings (Table S9). In a subsequent analyses we included the perception of men's intelligence in our model (Table S11).

In this second set of raters we replicated results of our main model: more intelligent men did not have a higher mate appeal. When analysing single ratings and not aggregated ratings, more intelligent men were even rated as less attractive as a short-term mate. In sum, we found no support for our hypothesis stating that more intelligent men have a higher mate appeal. Only men who were perceived to be more intelligent and men who were physically more attractive had a higher mate appeal.

Table S8.
Results of linear model predicting aggregated short-term and long-term mate attractiveness

	Short-term mate attractiveness			Long-term mate attractiveness		
Term	Estimate	p	95 % CI	Estimate	p	95% CI
Intercept	-0.73	.002	[-1.18; -0.28]	.05	.818	[-0.42; 0.53]
g factor	0.01	.902	[-0.11; 0.13]	.10	.121	[-0.03; 0.23]
physical attractiveness	1.11	<.001	[0.96; 1.27]	.92	<.001	[0.76; 1.08]

Note. 88 Targets were rated by n = 30 women rating men's short-term/long-term mate attractiveness and n = 19 women rating men's physical attractiveness.

Table S9.
Results of linear model predicting single ratings of short-term and long-term mate attractiveness

Term	Short-term mate attractiveness			Long-term mate attractiveness		
	Estimate	*p*	95 % CI	Estimate	*p*	95% CI
Intercept	1.62	<.001	[1.28; 1.96]	1.99	<.001	[1.64; 2.34]
g factor	-0.16	=.038	[-0.32; -0.01]	-0.04	.581	[-0.19; 0.10]
physical attractiveness	0.28	<.001	[0.20; 0.36]	0.23	<.001	[0.16; 0.30]
Observations	*n* = 49,316 ratings			*n* = 49,316 ratings		

Note. 88 target men were rated by *n* = 30 raters of short-term/long-term mate attractiveness and *n* = 19 raters of physical attractiveness.

Table S10.
Results of linear model predicting single ratings of short-term and long-term mate attractiveness without covariates

Term	Short-term mate attractiveness			Long-term mate attractiveness		
	Estimate	*p*	95% CI	Estimate	*p*	95% CI
Intercept	2.41	<.001	[2.34; 2.47]	2.64	<.001	[2.58; 2.70]
g factor	-0.22	<.001	[-0.29; -0.16]	-0.09	=.008	[-0.15; -0.02]
Observations	*n* = 2,620 ratings			*n* = 2,618 ratings		

Note. 88 target men were rated by n = 30 raters of short-term/long-term mate attractiveness.

Table S11
Results of linear model predicting single ratings of short-term and long-term mate attractiveness with perceived intelligence

Term	Short-term mate attractiveness			Long-term mate attractiveness		
	Estimate	*p*	95 % CI	Estimate	*p*	95% CI
Intercept	1.04	<.001	[0.60; 1.48]	1.12	<.001	[0.73; 1.51]
g factor	-0.19	=.013	[-0.34; -0.04]	-0.08	=.210	[-0.21; 0.05]
Perceived intelligence	0.17	=.002	[0.06; 0.29]	0.26	<.001	[0.17; 0.36]
physical attractiveness	0.27	<.001	[0.20; 0.35]	0.22	<.001	[0.16; 0.29]
Observations	*n* = 767,989 ratings			*n* = 767,400 ratings		

Note. 88 targets were rated by *n* = 30 raters of short-term/long-term mate attractiveness, *n* = 19 raters of physical attractiveness and *n* = 16 raters on perceived intelligence.

We assumed that above the influence of intelligence, funnier men have a higher mate appeal. We predicted men's short-term and long-term mate attractiveness, respectively, with men's *g* factor and ratings of their funniness. Replicating our previous results, men's actual intelligence did not influence their mate appeal. Again, when analysing single ratings more intelligent men were rated as less attractive as a short-term mate. However, men who were perceived to be funnier had a higher mate appeal (Table S12, Table S13). Despite the significant effect of funniness, results do not support our prediction. Perceived funniness contributes to mate appeal independently of men's intelligence. Hence, funniness does not seem to be an indicator of intelligence.

Table S12.
Results of linear model predicting aggregated short-term and long-term mate attractiveness with funniness

Term	Short-term mate attractiveness			Long-term mate attractiveness		
	Estimate	p	95 % CI	Estimate	*p*	95% CI
Intercept	-1.59	<.001	[-2.07; 1.10]	-1.02	<.001	[-1.48; -0.57]
g factor	0.00	.938	[-0.10; 0.11]	0.10	.056	[0.00; 0.19]
Perceived funniness	0.55	<.001	[0.36; 0.74]	0.69	<.001	[0.51; 0.87]
Physical attractiveness	0.84	<.001	[0.68; 1.00]	0.58	<.001	[0.42; 0.73]

Note. 88 Targets were rated by *n* = 30 women rating men's short-term/long-term mate attractiveness, *n* = 16 women rated men's funniness and *n* = 19 women rating men's physical attractiveness.

Table S13
Results of linear model predicting single ratings of short-term and long-term mate attractiveness with funniness

Term	Short-term mate attractiveness			Long-term mate attractiveness		
	Estimate	p	95 % CI	Estimate	p	95% CI
Intercept	0.69	<.001	[0.35; 1.03]	1.06	<.001	[0.68; 1.44]
g factor	-0.14	.034	[-0.26; -0.01]	-0.02	.792	[-0.13; 0.10]
Perceived funniness	0.35	<.001	[0.26; 0.45]	0.35	<.001	[0.26; 0.44]
Physical attractiveness	0.24	<.001	[0.17; 0.30]	0.19	<.001	[0.13; 0.24]
Observations	*n* = 767,989 ratings			*n* = 767,400 ratings		

Note. 88 target men were rated by *n* = 30 raters of short-term/long-term mate attractiveness, *n* = 19 raters of physical attractiveness and *n* = 16 raters of funniness.

And lastly, we assumed that when shifting from only physical attractiveness information (cue 1 -3), to additional cues related to men's intelligence (cue 4), short-term mate and long-term mate attractiveness of more intelligent men would increase. We predicted

a further increase for more intelligent men after presenting additional intelligence information (cue 6: make experimenter laugh).

Similar to our results on short-term mate attractiveness, we found none of the predicted effects for men's long-term mate attractiveness: attractiveness ratings for more intelligent men increased neither after cue 4 (*g* factor x Cue 4: *b* = .02, 95% HDI [-.01; .06]) nor after cue 6 (*g* factor x Cue 4: b = .04, 95% HDI [-.00; .08]).

Table S14

Estimated associations from a Bayesian mixed effect location scale model.

	Estimated effect on each outcome [95% HDI]			
	Short-term mate attractiveness		Long-term mate attractiveness	
Term	**Estimate**	**95% CI**	**Estimate**	**95% CI**
non-varying				
Intercept	+0.44	+0.02;+0.87	+0.44	-0.01;+0.87
σ Intercept	-0.03	-0.18;+0.13	-0.24	-0.43;-0.05
Study 1.4	-0.50	-1.04;+0.05	-0.59	-1.16;+0.00
Study 1.5	-0.71	-1.20;-0.19	-0.65	-1.16;-0.11
Cue 1-2	-0.30	-0.41;-0.20	-0.27	-0.37;-0.17
Cue 4	+0.16	+0.07;+0.24	+0.11	+0.04;+0.20
Cue 6	+0.36	+0.23;+0.50	+0.26	+0.13;+0.38
phys. attractiveness	+1.15	+1.05;+1.24	+1.01	+0.93;+1.09
g factor	-0.07	-0.11;-0.03	-0.02	-0.06;+0.01
Study 1.4:Cue 1-2	+0.26	+0.12;+0.39	+0.23	+0.09;+0.36
Study 1.5:Cue 1-2	+0.25	+0.11;+0.39	+0.21	+0.08;+0.35
Study 1.4:Cue 4	-0.05	-0.16;+0.04	-0.04	-0.14;+0.06
Study 1.5:Cue 4	-0.05	-0.16;+0.05	+0.00	-0.10;+0.11
Study 1.4:Cue 6	-0.19	-0.36;-0.03	-0.15	-0.31;+0.00
Study 1.5:Cue 6	-0.26	-0.43;-0.09	-0.17	-0.32;-0.02
Cue 1-2:phys. attractiveness	-0.10	-0.14;-0.06	-0.06	-0.10;-0.03
Cue 4:phys. attractiveness	+0.07	+0.04;+0.11	+0.06	+0.02;+0.10
Cue 6:phys. attractiveness	+0.14	+0.10;+0.19	+0.13	+0.09;+0.18
Cue 1-2:*g* factor	-0.01	-0.04;+0.02	-0.01	-0.04;+0.01
Cue 4:*g* factor	+0.01	-0.02;+0.04	+0.02	-0.01;+0.06
Cue 6:*g* factor	+0.02	-0.02;+0.06	+0.04	-0.00;+0.08
σ Study 1.4	+0.28	+0.08;+0.47	+0.55	+0.31;+0.81
σ Study 1.5	+0.29	+0.11;+0.47	+0.47	+0.24;+0.69
σ Cue 1-2	-0.11	-0.15;-0.08	-0.14	-0.17;-0.10
σ Cue 4	+0.13	+0.08;+0.17	+0.16	+0.12;+0.21
σ Cue 6	+0.29	+0.24;+0.33	+0.34	+0.29;+0.38
σ Study 1.4:Cue 1-2	+0.12	+0.07;+0.17	+0.14	+0.09;+0.19
σ Study 1.5:Cue 1-2	+0.09	+0.03;+0.14	+0.11	+0.06;+0.17
σ Study 1.4:Cue 4	-0.04	-0.10;+0.02	-0.07	-0.13;-0.01
σ Study 1.5:Cue 4	-0.04	-0.10;+0.01	-0.05	-0.11;+0.00
σ Study 1.4:Cue 6	-0.09	-0.15;-0.03	-0.14	-0.20;-0.08
σ Study 1.5:Cue 6	-0.05	-0.11;+0.01	-0.05	-0.11;+0.01
rater (n=114)				
sd(Intercept)	+1.13	+0.98;+1.30	+1.17	+1.02;+1.34
sd(phys. attractiveness)	+0.46	+0.40;+0.53	+0.44	+0.38;+0.51

sd(*g* factor)	+0.16	+0.13;+0.20	+0.15	+0.12;+0.18
sd(Cue 1-2)	+0.21	+0.17;+0.26	+0.21	+0.16;+0.26
sd(Cue 4)	+0.11	+0.07;+0.15	+0.10	+0.05;+0.15
sd(Cue 6)	+0.26	+0.21;+0.32	+0.24	+0.18;+0.30
sd(phys. attractiveness:Cue 1-2)	+0.06	+0.03;+0.09	+0.03	+0.01;+0.06
sd(phys. attractiveness:Cue 4)	+0.02	+0.00;+0.06	+0.03	+0.01;+0.06
sd(phys. attractiveness:Cue 6)	+0.07	+0.03;+0.11	+0.09	+0.05;+0.14
sd(*g* factor:Cue 1-2)	+0.02	+0.00;+0.05	+0.01	+0.00;+0.04
sd(*g* factor:Cue 4)	+0.01	+0.00;+0.04	+0.03	+0.00;+0.06
sd(*g* factor:Cue 6)	+0.02	+0.00;+0.06	+0.04	+0.00;+0.08
sd(σ Intercept)	+0.38	+0.33;+0.43	+0.48	+0.42;+0.55
target (n=88)				
sd(Intercept)	+0.19	+0.16;+0.24	+0.08	+0.06;+0.10
sd(Cue 1-2)	+0.07	+0.04;+0.11	+0.03	+0.01;+0.05
sd(Cue 4)	+0.08	+0.04;+0.12	+0.03	+0.00;+0.05
sd(Cue 6)	+0.14	+0.09;+0.19	+0.12	+0.08;+0.15
sd(σ Intercept)	+0.28	+0.24;+0.33	+0.28	+0.24;+0.33

Note:

Estimated associations from a Bayesian mixed effect location scale model. Estimates prefixed σ denote estimates on the residual standard deviation.

C. Robustness Checks: Inclusion of only heterosexual targets and raters

We pre-registered to recruit only heterosexual participants. However, three of our target men reported a bisexual or homosexual orientation. 12 of our female raters reported a bisexual or homosexual orientation. We excluded these 15 participants and reran our analyses. Replicating our results, more intelligent men were not preferred as a potential partner (for short-term mate attractiveness $b = -.07$, 95% HDI [-.11;-.03], for long-term mate attractiveness $b = -.03$, 95% HDI [-.06, .01]). Additionally, attractiveness did not increase after increasing information on men's intelligence (for short-term mate attractiveness: cue 4 * g factor $b = .01$, 95% HDI [-.02, .05], cue 6 * g factor $b = .03$, 95% HDI [-.01, .07]; for long-term mate attractiveness: cue 4 * g factor $b = .03$, 95% HDI [-.01, .07], cue 6 * g factor $b = -.04$, 95% HDI [-.00, .09]).

Table S15

Estimated associations from a Bayesian mixed effect location scale model including only heterosexual participants

	Estimated effect on each outcome [95% HDI]			
	Short-term mate attractiveness		**Long-term mate attractiveness**	
Term	**Estimate**	**95% CI**	**Estimate**	**95% CI**
non-varying				
Intercept	+0.41	-0.02;+0.85	+0.42	-0.04;+0.89
σ Intercept	-0.04	-0.20;+0.13	-0.24	-0.46;-0.04
Study 1.4	-0.53	-1.10;+0.02	-0.61	-1.24;+0.01
Study 1.5	-0.66	-1.17;-0.14	-0.60	-1.17;-0.03
Cue 1-2	-0.30	-0.40;-0.19	-0.26	-0.37;-0.16
Cue 4	+0.16	+0.08;+0.24	+0.11	+0.04;+0.19
Cue 6	+0.36	+0.23;+0.49	+0.25	+0.12;+0.37
phys. attractiveness	+1.11	+1.02;+1.20	+0.98	+0.89;+1.06
g factor	-0.07	-0.11;-0.03	-0.03	-0.06;+0.01
Study 1.4:Cue 1-2	+0.27	+0.14;+0.41	+0.24	+0.09;+0.38
Study 1.5:Cue 1-2	+0.23	+0.09;+0.37	+0.20	+0.06;+0.34
Study 1.4:Cue 4	-0.06	-0.16;+0.04	-0.04	-0.15;+0.06
Study 1.5:Cue 4	-0.03	-0.14;+0.08	+0.03	-0.08;+0.13
Study 1.4:Cue 6	-0.20	-0.37;-0.04	-0.15	-0.31;+0.00
Study 1.5:Cue 6	-0.22	-0.38;-0.05	-0.13	-0.28;+0.03
Cue 1-2:phys. attractiveness	-0.10	-0.13;-0.06	-0.06	-0.10;-0.03
Cue 4:phys. attractiveness	+0.07	+0.04;+0.11	+0.06	+0.02;+0.10
Cue 6:phys. attractiveness	+0.15	+0.10;+0.19	+0.14	+0.09;+0.18
Cue 1-2:g factor	-0.01	-0.04;+0.02	-0.01	-0.04;+0.01
Cue 4:g factor	+0.01	-0.02;+0.05	+0.03	-0.01;+0.07
Cue 6:g factor	+0.03	-0.01;+0.07	+0.04	-0.00;+0.09
σ Study 1.4	+0.28	+0.06;+0.49	+0.54	+0.28;+0.81
σ Study 1.5	+0.30	+0.10;+0.48	+0.47	+0.23;+0.72
σ Cue 1-2	-0.11	-0.15;-0.08	-0.14	-0.17;-0.10
σ Cue 4	+0.13	+0.08;+0.17	+0.17	+0.12;+0.21
σ Cue 6	+0.29	+0.24;+0.33	+0.34	+0.29;+0.38
σ Study 1.4:Cue 1-2	+0.12	+0.07;+0.17	+0.14	+0.09;+0.20
σ Study 1.5:Cue 1-2	+0.09	+0.03;+0.15	+0.12	+0.06;+0.18

σ Study 1.4:Cue 4	-0.04	-0.10;+0.02	-0.07	-0.13;-0.01
σ Study 1.5:Cue 4	-0.04	-0.10;+0.02	-0.06	-0.12;+0.00
σ Study 1.4:Cue 6	-0.10	-0.16;-0.04	-0.15	-0.21;-0.09
σ Study 1.5:Cue 6	-0.06	-0.12;+0.01	-0.05	-0.11;+0.01
rater (n=105)				
sd(Intercept)	+1.16	+1.01;+1.34	+1.21	+1.05;+1.39
sd(phys. attractiveness)	+0.45	+0.39;+0.52	+0.43	+0.37;+0.50
sd(g factor)	+0.16	+0.12;+0.19	+0.14	+0.11;+0.18
sd(Cue 1-2)	+0.21	+0.17;+0.27	+0.21	+0.16;+0.27
sd(Cue 4)	+0.10	+0.07;+0.14	+0.09	+0.04;+0.14
sd(Cue 6)	+0.26	+0.21;+0.32	+0.23	+0.17;+0.30
sd(phys. attractiveness:Cue 1-2)	+0.06	+0.03;+0.10	+0.03	+0.01;+0.07
sd(phys. attractiveness:Cue 4)	+0.02	+0.00;+0.06	+0.03	+0.01;+0.06
sd(phys. attractiveness:Cue 6)	+0.07	+0.03;+0.11	+0.09	+0.04;+0.14
sd(g factor:Cue 1-2)	+0.02	+0.00;+0.05	+0.01	+0.00;+0.04
sd(g factor:Cue 4)	+0.01	+0.00;+0.04	+0.03	+0.00;+0.06
sd(g factor:Cue 6)	+0.02	+0.00;+0.06	+0.04	+0.00;+0.09
sd(σ Intercept)	+0.39	+0.34;+0.45	+0.50	+0.44;+0.58
target (n=88)				
sd(Intercept)	+0.19	+0.16;+0.23	+0.08	+0.05;+0.10
sd(Cue 1-2)	+0.07	+0.03;+0.10	+0.03	+0.00;+0.05
sd(Cue 4)	+0.07	+0.03;+0.11	+0.02	+0.00;+0.05
sd(Cue 6)	+0.13	+0.08;+0.19	+0.12	+0.08;+0.15
sd(σ Intercept)	+0.29	+0.25;+0.34	+0.29	+0.25;+0.34

Note:

Estimated associations from a Bayesian mixed effect location scale model. Estimates prefixed σ denote estimates on the residual standard deviation.

D. Robustness Check: No effects on sigma

Table S16.
Results for a simple mixed effects model

	Estimated effect on each outcome [95% CI]			
	Short-term mate attractiveness		Long-term mate attractiveness	
Term	**Estimate**	**95% CI**	**Estimate**	**95% CI**
non-varying				
Intercept	+0.41	-0.04;+0.86	+0.39	-0.08;+0.86
Study 1.4	-0.49	-1.08;+0.09	-0.53	-1.11;+0.09
Study 1.5	-0.60	-1.14;-0.07	-0.53	-1.07;+0.04
Cue 1-2	-0.38	-0.51;-0.24	-0.34	-0.46;-0.21
Cue 4	+0.12	+0.01;+0.23	+0.11	-0.01;+0.23
Cue 6	+0.33	+0.18;+0.49	+0.26	+0.09;+0.43
phys. attractiveness	+1.20	+1.11;+1.29	+1.06	+0.97;+1.15
g factor	-0.09	-0.15;-0.04	-0.05	-0.10;+0.00
Study 1.4:Cue 1-2	+0.32	+0.14;+0.49	+0.29	+0.12;+0.46
Study 1.5:Cue 1-2	+0.34	+0.17;+0.51	+0.29	+0.12;+0.45
Study 1.4:Cue 4	-0.03	-0.17;+0.12	-0.03	-0.18;+0.12
Study 1.5:Cue 4	+0.01	-0.13;+0.15	+0.02	-0.13;+0.17
Study 1.4:Cue 6	-0.18	-0.39;+0.01	-0.20	-0.41;+0.01
Study 1.5:Cue 6	-0.20	-0.39;-0.00	-0.14	-0.33;+0.05
Cue 1-2:phys. attractiveness	-0.14	-0.19;-0.09	-0.10	-0.15;-0.05
Cue 4:phys. attractiveness	+0.09	+0.04;+0.15	+0.06	+0.01;+0.12
Cue 6:phys. attractiveness	+0.18	+0.12;+0.23	+0.11	+0.05;+0.16
Cue 1-2:g factor	-0.01	-0.06;+0.04	-0.02	-0.07;+0.03
Cue 4:g factor	+0.04	-0.01;+0.09	+0.08	+0.03;+0.13
Cue 6:g factor	+0.05	-0.00;+0.11	+0.09	+0.04;+0.15
rater (n=114)				
sd(Intercept)	+1.12	+0.98;+1.30	+1.16	+1.01;+1.33
sd(phys. attractiveness)	+0.45	+0.39;+0.52	+0.44	+0.38;+0.51
sd(g factor)	+0.19	+0.16;+0.22	+0.17	+0.14;+0.21
sd(Cue 1-2)	+0.24	+0.18;+0.30	+0.24	+0.18;+0.30
sd(Cue 4)	+0.14	+0.08;+0.19	+0.16	+0.10;+0.21
sd(Cue 6)	+0.28	+0.22;+0.35	+0.30	+0.24;+0.37
sd(phys. attractiveness:Cue 1-2)	+0.07	+0.01;+0.12	+0.04	+0.00;+0.09
sd(phys. attractiveness:Cue 4)	+0.02	+0.00;+0.06	+0.02	+0.00;+0.06
sd(phys. attractiveness:Cue 6)	+0.04	+0.00;+0.10	+0.06	+0.00;+0.12
sd(g factor:Cue 1-2)	+0.02	+0.00;+0.06	+0.03	+0.00;+0.08
sd(g factor:Cue 4)	+0.02	+0.00;+0.06	+0.04	+0.00;+0.09
sd(g factor:Cue 6)	+0.03	+0.00;+0.09	+0.07	+0.01;+0.13
target (n=88)				
sd(Intercept)	+0.31	+0.26;+0.38	+0.26	+0.21;+0.31
sd(Cue 1-2)	+0.15	+0.10;+0.20	+0.13	+0.08;+0.19
sd(Cue 4)	+0.13	+0.08;+0.19	+0.19	+0.13;+0.25
sd(Cue 6)	+0.23	+0.17;+0.29	+0.28	+0.22;+0.35

Note:
Estimated associations from a Bayesian mixed effects model without allowing the residual variation to vary. In this model, the 95% HDI for the interactions between Cue 4/6 and the g factor on long-term mate attractiveness excluded zero, but this result was not robust in models that allowed residual variation to vary across rating studies.

Fig. S1: The aggregated long-term mate attractiveness ratings made after seeing each cue (or set of cues) was adjusted for physical attractiveness. The plot shows the slope of a linear regression predicting short-term attractiveness from the measured *g* factor. More intelligent men were rated only slightly more favourably after intelligence-relevant information became available.
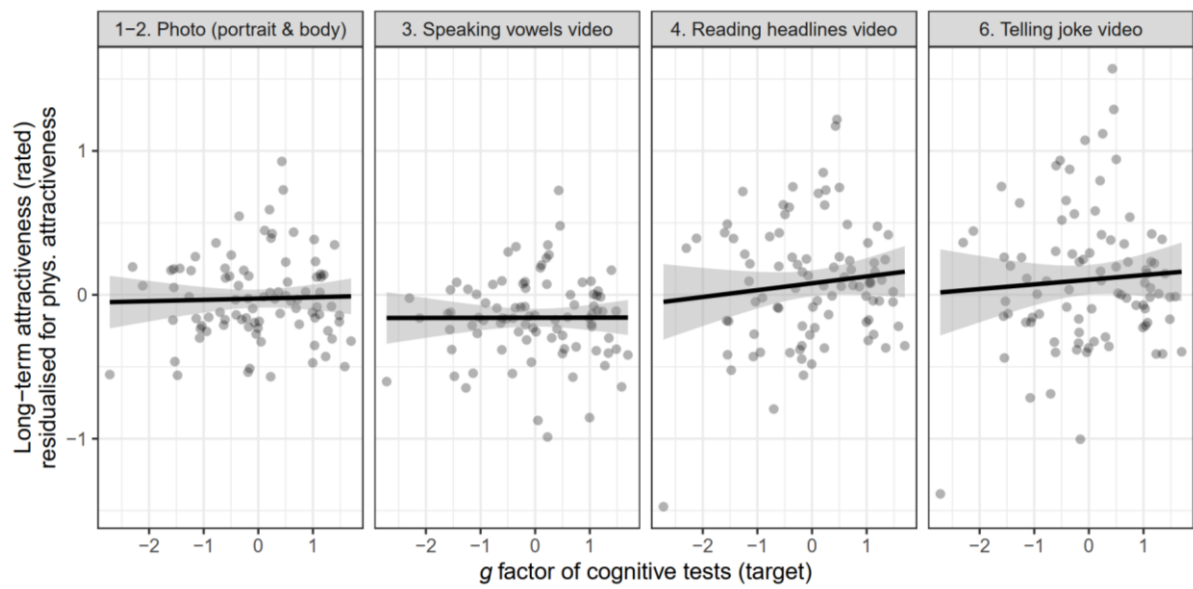
E. Results remain unchanged with known participants and partnered participants excluded

Table S17.
Study 2 results excluding known participants and partnered participants

Predictors	Rated Intelligence			Rated Funniness			Rated Attractiveness		
	Estimates	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
Intercept	-0.06	-0.14 – 0.02	0.150	0.08	-0.00 – 0.16	0.063	0.09	-0.00 – 0.18	0.053
Shipley (Vocabulary)	0.07	0.02 – 0.12	**0.008**	0.01	-0.05 – 0.06	0.834	-0.01	-0.06 – 0.05	0.811
Sex (Female)	0.13	0.01 – 0.25	**0.028**	-0.13	-0.25 – -0.01	**0.031**	-0.15	-0.28 – -0.02	**0.019**
N	671 $_{id}$			671 $_{id}$			671 $_{id}$		
	740 $_{partnerid}$			740 $_{partnerid}$			740 $_{partnerid}$		
Observations	1977			1977			1975		

Predictors	Rated Attractiveness			Rated Funniness			Rated Attractiveness		
	Estimates	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
Intercept	0.12	0.03 – 0.20	**0.005**	0.09	0.01 – 0.17	**0.020**	0.06	-0.01 – 0.14	0.109
Rated Intelligence	0.29	0.25 – 0.33	**<0.001**	0.28	0.23 – 0.32	**<0.001**			
Sex (Female)	-0.20	-0.32 – -0.09	**0.001**	-0.17	-0.28 – -0.06	**0.003**	-0.11	-0.22 – -0.00	**0.043**
Rated Funniness							0.41	0.37 – 0.44	**<0.001**
N	695 $_{id}$			695 $_{id}$			695 $_{id}$		
	743 $_{partnerid}$			743 $_{partnerid}$			743 $_{partnerid}$		
Observations	2045			2047			2045		

F. No moderation by sex

Table S18.
Study 2 results including moderation by sex and all random effects.

Predictors	Rated Intelligence			Rated Funniness			Rated Attractiveness		
	Estimates	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
Intercept	-0.07	-0.15 – 0.01	0.091	0.07	-0.01 – 0.15	0.098	0.08	-0.01 – 0.16	0.076
Shipley (Vocabulary)	0.08	0.01 – 0.15	**0.026**	-0.01	-0.09 – 0.06	0.720	0.01	-0.07 – 0.09	0.780
Sex (Female)	0.14	0.02 – 0.25	**0.017**	-0.13	-0.25 – -0.02	**0.026**	-0.13	-0.25 – -0.01	**0.030**
Shipley (Vocabulary):Sex	-0.00	-0.10 – 0.10	0.998	0.01	-0.09 – 0.11	0.876	-0.05	-0.16 – 0.06	0.376
Random Effects									
σ^2	0.58			0.56			0.47		
τ_{00}	0.24 partnerid			0.22 partnerid			0.22 partnerid		
	0.17 id			0.22 id			0.29 id		
τ_{11}	0.01 partnerid.scale(shipleys)			0.01 partnerid.scale(shipleys)			0.01 partnerid.scale(shipleys)		
ρ_{01}	0.07 partnerid			0.00 partnerid			0.19 partnerid		
ICC	0.42			0.44			0.53		
N	729 id			729 id			729 id		
	757 partnerid			757 partnerid			757 partnerid		
Observations	2247			2247			2245		

	Rated Attractiveness			Rated Funniness			Rated Attractiveness		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
Intercept	0.12	0.04 – 0.19	**0.004**	0.09	0.02 – 0.17	**0.012**	0.06	-0.02 – 0.13	0.138
Rated Intelligence	0.34	0.28 – 0.40	**<0.001**	0.33	0.27 – 0.39	**<0.001**			
Sex	-0.20	-0.31 – -0.09	**<0.001**	-0.19	-0.29 – -0.08	**<0.001**	-0.09	-0.20 – 0.01	0.075
Rated Intelligence:Sex	-0.07	-0.16 – 0.01	0.087	-0.07	-0.15 – 0.01	0.105			
Rated Funniness							0.38	0.33 – 0.43	**<0.001**
Rated Funniness:Sex							0.06	-0.01 – 0.13	0.116

Random Effects

σ^2	0.39			0.50			0.42		
τ_{00}	0.15 $_{partnerid}$			0.13 $_{partnerid}$			0.16 $_{partnerid}$		
	0.29 $_{id}$			0.21 $_{id}$			0.22 $_{id}$		
τ_{11}	0.05 $_{partnerid.scale(rintelligent)}$			0.02 $_{partnerid.scale(rintelligent)}$					
	0.01 $_{id.scale(rintelligent)}$			0.04 $_{id.scale(rintelligent)}$					
ρ_{01}	0.26 $_{partnerid}$			0.44 $_{partnerid}$					
	0.14 $_{id}$			0.23 $_{id}$					
ICC	0.56			0.44			0.47		
N	753 $_{id}$			753 $_{id}$			753 $_{id}$		
	758 $_{partnerid}$			758 $_{partnerid}$			758 $_{partnerid}$		
Observations	2319			2321			2319		

G. Pattern of results remains the same when facial and bodily attractiveness is controlled

Table S19.
Study 2 results with facial and bodily attractiveness variables included in all models

Predictors	Rated Intelligence			Rated Funniness			Rated Overall Attractiveness		
	Estimates	CI	p	Estimates	CI	p	Estimates	CI	p
(Intercept)	-0.09	-0.16 – -0.01	**0.021**	0.04	-0.04 – 0.11	0.319	0.00	-0.04 – 0.04	0.978
Shipley (Vocabulary)	0.08	0.03 – 0.13	**0.001**	-0.00	-0.05 – 0.04	0.864	-0.01	-0.03 – 0.02	0.672
Facial Attractiveness	0.19	0.13 – 0.24	**<0.001**	0.27	0.22 – 0.33	**<0.001**	0.59	0.56 – 0.62	**<0.001**
Body Attractiveness	0.11	0.06 – 0.17	**<0.001**	0.15	0.09 – 0.20	**<0.001**	0.29	0.26 – 0.32	**<0.001**
Sex (Female)	0.18	0.07 – 0.28	**0.001**	-0.07	-0.18 – 0.03	0.169	-0.00	-0.06 – 0.05	0.939
N	729 id			729 id			729 id		
	757 partnerid			757 partnerid			757 partnerid		
Observations	2246			2246			2244		

Predictors	Rated Overall Attractiveness			Rated Funniness			Rated Overall Attractiveness		
	Estimates	CI	p	Estimates	CI	p	Estimates	CI	p
(Intercept)	0.01	-0.02 – 0.05	0.485	0.06	-0.01 – 0.13	0.104	-0.00	-0.04 – 0.04	0.951
Rated Intelligence	0.09	0.06 – 0.11	**<0.001**	0.20	0.17 – 0.24	**<0.001**			
Facial Attractiveness	0.57	0.54 – 0.61	**<0.001**	0.24	0.19 – 0.29	**<0.001**	0.55	0.52 – 0.58	**<0.001**
Bodily Attractiveness	0.28	0.25 – 0.31	**<0.001**	0.12	0.07 – 0.18	**<0.001**	0.27	0.24 – 0.30	**<0.001**
Sex (Female)	-0.02	-0.08 – 0.03	0.389	-0.11	-0.21 – -0.02	**0.024**	0.00	-0.05 – 0.06	0.853
Rated Funniness							0.14	0.12 – 0.16	**<0.001**
N	753 id			753 id			753 id		
	758 partnerid			758 partnerid			758 partnerid		
Observations	2317			2319			2317		

Appendix B.

Manuscript 2

Driebe, J.C., Stern, J., Penke, L. & Gerlach, T.M. (2021). *Ideal Partner Preferences are Stable Over 13 Years* [Manuscript submitted for publication]. Department of Biological Personality Psychology, University of Goettingen.

This article has been submitted for publication in the Journal Social Psychological and Personality Science (SPPS), published by Sage Journal.

Ideal Partner Preferences are Stable Over 13 Years

Julie C. Driebe[1], Julia Stern[1,3], Lars Penke[1,2] & Tanja M. Gerlach[1,2]

[1]Department of Psychology, University of Goettingen

[2]Leibniz ScienceCampus Primate Cognition, Goettingen

[3]Department of Psychology, University of Bremen

corresponding author: juliedriebe@gmail.com

Julie C. Driebe is a doctoral candidate at the University of Goettingen. Her research focuses on partner preferences and their role in romantic relationships.

Julia Stern is a researcher at the Universities of Bremen and Goettingen. Among other interests, her research focuses on ovulatory cycle effects, personality and romantic relationships.

Lars Penke is a Professor of Psychology at the University of Goettingen. He is interested in the biological foundations of individual differences, including evolutionary perspectives on mate choice, romantic relationships, and sexuality.

Tanja M. Gerlach is a senior researcher at the University of Goettingen. Her research focuses on how people establish and maintain romantic relationships and the role individual differences, such as narcissism and romantic ideals, play therein.

Word count: 5000

Abstract

Ideal partner preferences are said to be stable cognitive constructs guiding relationship decisions. However, no study investigated a time span exceeding three years. Here, we investigate stability and change of partner preferences across a much longer time span of 13 years. In addition, we investigate participants' insight into how their preferences have changed. 204 participants ($M = 46.2$ years, $SD = 7.4$, 104 women) reported their partner preferences at two time points. We found a mean rank-order stability of $r = .42$ and an overall profile stability of $r = .73$ (distinctive $r = .53$). Some preferences changed over time, e.g. increased for status-resources, which was related to age and parenthood. Finally, we found some, but varying insight into how preferences had changed (mean $r = .20$). Results support the idea of partner preferences being stable cognitive constructs, but suggest some variability related to demands of different life stages.

Keywords: ideal partner preferences, stability, change, perceived change

Thinking back to our expectations of an ideal partner a decade ago and comparing these expectations between now and then might lead to the feeling that our ideas about Mr. or Mrs. Right have changed tremendously. But is this really the case? With the current study we aim to answer this question: We investigate the stability and change of ideal partner preferences across a time span of 13 years and explore whether individuals possess insight into how their preferences have changed.

Ideal Partner Preferences

With ideal partner preferences, we refer to the aspirations or standards an individual has about an ideal partner (Simpson et al., 2001). These standards are used to evaluate (potential) partners and should thereby guide relationship decisions (Fletcher et al., 1999; Simpson et al., 2001). The Ideal Standards Model provides a framework to describe the qualities of an ideal partner using three correlated factors: Warmth-trustworthiness (WT), vitality-attractiveness (VA) and status-resources (SR) (Fletcher et al., 1999). These three factors are well-replicated (Campbell & Fletcher, 2015). However, as suggested by other studies (e.g. Shackelford et al., 2005b), qualities describing an ideal partner may not be limited to these three factors: Other factors that have been reported include, for example, family orientation (Lam et al., 2016); intellect, dominance (Csajbók & Berkics, 2017; Schwarz & Hassebrauck, 2012); or humour and sociability (Schwarz & Hassebrauck, 2012).

Although Fletcher and colleagues (1999) defined ideal partner preferences as stable cognitive constructs which differ between individuals, evidence on the stability of these preferences is scarce. The largest retest stability so far was found after a period of three weeks $r = .85$[6] (Fletcher et al., 1999) and three months $r = .82$[1] (Fletcher et al., 2000). Retest correlations seem to reduce substantially as more time goes by, e.g. to $r = .65$ after five months (Gerlach, Arslan, et al., 2019), $r = .55$[1] after 27 months (Bredow & Hames, 2019), and $r = .51$[1] (Shackelford et al., 2005a) or $r = .35$ after three years (Bleske-Rechek & Ryan, 2015). But what happens after a longer time span? If ideal partner preferences are stable, the strongest declines in stability coefficients may be found in the first years after assessment and stabilise thereafter, similar to stability in personality traits (Anusic & Schimmack, 2016; Costa et al., 2019). To our knowledge, so far, no study investigated how stable ideals are for a

---

[6] We calculated mean-retest correlations by calculating the mean of Fisher z-transformed correlation coefficients as they were originally reported by Fletcher and colleagues (1999), Fletcher and colleagues (2000), Bredow and Hames (2019) and Shackelford and colleagues (2005a).

time span exceeding three years. If these standards used to evaluate partners are indeed stable constructs, then they should show substantial stability even over an extended period of time.

However, it is also plausible that ideals change across a longer time span, hence, that the stability decreases after more and more years. For example, developmental theories on human motivation posit that individuals face different challenges during life that go along with a shift in priorities (Heckhausen et al., 2010, 2019). Similarly, Life-History Theory (LHT) may give a rationale for changes in partner preferences. LHT proposes that every individual has a limited budget of effort and resources (Alexander, 1987; Del Giudice et al., 2016). Across the life-span, individuals should therefore face trade-offs in what kind of activities they allocate their energy to, with one trade-off existing between parental and mating effort (Del Giudice et al., 2016). Accordingly, individuals could experience shifts in their ideals related to the demands of different life stages. For example, a preference for attractiveness could especially be important during life stages where much effort is allocated to mating and finding a young and healthy partner could increase the offspring's quality (Buss & Schmitt, 1993; Fletcher et al., 1999). A preference for resources could be especially important a bit later in life, during stages where more effort is allocated into parenting, when more resources are needed to provide for offspring (Campbell & Fletcher, 2015). Because secure relationships characterised by high levels of cooperation and support should always be beneficial, a preference for WT might be less susceptible to change. Consequently, as individuals face different challenges during life going along with a shift in priorities, ideals might change across the lifespan.

Insight into Changes of Ideal Partner Preferences

People tend to have opinions on whether they prefer the same type of partner as they already did a decade ago, or on how much their preferences have changed. But are these opinions accurate? Two previous studies addressed perceptions of change in ideal partner preferences. Sprecher and colleagues (2018) asked participants how they believed to have changed across two to three years. Participants believed to place higher importance on all dimensions, whereas older individuals assumed to have changed to a smaller degree than younger individuals. Bleske-Rechek and colleagues (2009) asked students how they believe ideals would change during college. Participants expected a partner's personality to become more and appearance to become less important. In additional samples, the authors investigated whether these beliefs mapped onto differences in ideals across different age

groups and found that, corresponding to expected changes, a partner's personality became more and appearance became less important with increasing age. However, none of these studies directly investigated whether perceived changes correspond to actual changes in ideals. Given that such changes are an intra-individual process, a more direct approach to investigate insight into preference change would be a longitudinal design.

The Current Project

In the current study, we followed up upon individuals of a former study that assessed participants' ideal partner preferences (among other measures) in the year 2006. Approximately 13 years later, we assessed these ideals for a second time and investigated whether participants perceive to have changed their preferences.

As there are different approaches to investigate the stability of ideals that come with different benefits, we investigate multiple stability indices. Rank-order stability estimates the degree to which the relative position of each individual in a population remains the same across time (Roberts & DelVecchio, 2000), one trait at a time. In contrast, profile correlations inform about the stability of a person's whole trait profile. Finally, when investigating mean-level stability, the average change of a trait in a population is examined across time, which indicates the direction of change (Asendorpf & Neyer, 2012).

We expect that ideals are stable over 13 years and predict that initial ideal partner preferences correlate positively with current partner preferences (H1). Nevertheless, we expect varying stabilities across different preference dimensions to reflect the varying demands of different life stages: We predict that the average preference for status-resources (H2) and family orientation (H3) increases from T1 to T2, especially when participants were younger at T1 (H2.1, H3.1). Moreover, we predict that the desire for status-resources changes with the immediacy of a desire for or actual existence of children (H4). Hence, the correlation of initial with current ideal partner preferences for status-resources should be smaller in a subsample of participants who now have children or are currently planning to have children compared to participants without children. Finally, we predict that people's perceptions of change in their ideals correspond to their actual changes in preferences for status-resources (H5.1) and vitality-attractiveness (H5.2).

Method

Participants and Recruitment

We re-recruited participants of the Berlin Speed Dating Study [BSDS] that was conducted in 2006 (Asendorpf et al., 2011). This study consisted of 382 participants (age $M = 32.8$ years, $SD = 7.4$, range $= 18 – 54$, female $= 192$, male $= 190$) which we tried to contact after around 13 years.

At the initial assessment (T1), participants gave us detailed contact information. From February to November 2019, we reached out to those former participants for a re-assessment (T2). As an incentive, participants received individual feedback on their personality, a comparison of their initial and current ideals and 40€, with the chance to receive a bonus of 10€[7].

We were able to recruit 226 participants (41% dropout) of the BSDS. We excluded four participants because they reported a homosexual orientation at T2 and 18 participants who dropped out during the T2 assessment. Our final sample size comprised $N = 204$ participants (age $M = 46.2$ years, $SD = 7.4$, range $= 31 – 66$; female $= 104$, male $= 100$). 64% participants were currently in a relationship, 35% were single and less than 1% reported an undefined relationship status. 84% participants reported having a university degree. An attrition analysis revealed that participants who completed participation at T2 were less conscientious (Hedges $g = -0.20$) and more neurotic (Hedges $g = 0.23$) compared to participants who participated only at T1. We found no other systematic group differences in demographics and other personality traits (see S1).

Because we could not foresee how many people we would be able to re-recruit for T2, we invited participants of another earlier study, the Sociosexuality Study, taking place between 2004 and 2005 (Penke & Asendorpf, 2008a) in order to increase our chance of a sufficient sample size[8]. However, we decided to run our main analyses based on participants of the BSDS only and include analyses of this second sample in our supplement because this sample

---

[7] This bonus was disbursed if participants invited their peers and partners to a second part of the data collection.
[8] At T1, this sample consisted of $N = 142$ participants, of which $n = 66$ individuals could be re-recruited for a T2 re-assessment.

came with several major limitations[9] and was difficult to compare to the BSDS re-assessment, yet turned out to be too small to analyse on its own (see S1B, S3).

Procedure

The re-assessment was implemented on the survey framework formr.org (Arslan et al., 2020a). We invited participants to our online study with the aim to investigate their relationship history longitudinally. After participants agreed to participate, they were asked to fill out a short demographic questionnaire including questions on their age, sex, relationship status, and family situation (Table 1). In a second step, participants were asked to rate various items regarding their importance in an ideal partner. Step three and four are of minor importance for the current study (detailed assessment of their relationship history). In a fifth step, participants finished their participation by answering different measures on their personality including questions on how they perceive their ideals had changed since their initial participation. Although not part of the current study, participants were finally asked to invite peers and their current partner to participate in a second part of the re-assessment.

---

[9] Aside from the small sample size, this second sample came with the limitation of a far less extensive assessment of ideal partner preferences. Only 13 items assessed ideal partner preferences and each entailed several characteristics, e.g., "parental abilities, wish for children". These items and response formats differed from the BSDS.

Table 1

Item Content and Response Formats of Demographic Questionnaire used in this Study.

Item	Content [Response format]
Sex	Your biological sex: [female, male]
Age	Your current age: [number between 18 to 100]
Relationship Status	What is your relationship status? [Single, uncommitted relationship, committed relationship, engaged, married, other]
Children	Do you have children? [Yes, No]
No. of children[a]	How many biological children do you have? [number between 1 and 40]
Wish for children[b]	Do you ever want to have children? [Yes, No, Do not know yet]
Pregnant trying	Do you currently try to become a father/mother (again)? [1: try to avoid it - 5: trying]

Note. [a] = The item was only presented to participants who indicated to have children. [b] = The item was only presented to participants who indicated not to have children.

Measures

Ideal Partner Preferences

At T1 and T2, participants rated 58 characteristics[10] on their importance in an ideal partner on a scale ranging from 1 (*very unimportant)* to 5 (*very important).* In a principal component analysis with oblimin rotation from a pre-test (see S1C), we extracted eight factors, which we labelled warmth-trustworthiness (WT), vitality-attractiveness (VA), status-resources (SR), family orientation (FO), intelligence (IQ), creativity (C), humour (H), and adventurousness-confidence (AC) (Table 2).

---

[10] At T2, participants rated 72 items and 59 of these items at T1. Thus, we excluded 14 items: 13 items because of missing values at T1 and one item because it did not load on any factor.

Table 2

Cronbach's Alpha, Means and Standard Deviations of all Items Assessing Ideal Partner Preferences at T1 and T2

Item of each dimension	a_{T1}	M_{T1} (SD_{T1})	a_{T2}	M_{T2} (SD_{T2})
WT	0.82		0.90	
understanding		4.15 (0.67)		4.28 (0.78)
sensitive		4.22 (0.63)		4.27 (0.70)
trustworthy		4.45 (0.61)		4.52 (0.66)
good listener		4.07 (0.72)		4.19 (0.82)
honest		4.67 (0.55)		4.56 (0.67)
considerate		3.97 (0.68)		4.20 (0.77)
supportive		3.79 (0.75)		4.14 (0.83)
faithful		4.39(0.76)		4.22 (0.94)
kind		3.74 (0.96)		3.89 (0.99)
VA	0.80		0.85	
erotic		3.82 (0.72)		3.69 (0.87)
sexy		3.57 (0.91)		3.50 (0.89)
arousing		3.90 (0.77)		3.54 (0.88)
attractive		3.91 (0.74)		3.77 (0.81)
nice body		3.51 (0.85)		3.51 (0.90)
appealing		4.05 (0.68)		3.96 (0.74)
passionate		3.92 (0.76)		3.63 (0.88)
sporty and athletic		3.31 (0.93)		3.37 (0.91)
fit		3.49 (0.82)		3.62 (0.83)
feminine		2.88 (1.36)		2.87 (1.25)
SR	0.84		0.85	
prosperous		2.12 (0.94)		2.45 (0.95)
good job		2.92 (0.92)		3.10 (0.92)
financially secure		2.83 (1.05)		3.06 (1.07)
successful		2.81 (0.90)		2.97 (0.92)
influential		2.19 (0.90)		2.28 (0.91)
of good standing		2.45 (1.02)		2.55 (0.93)
good family background		2.01 (1.01)		2.14 (1.05)
nice house or apartment		2.67 (1.06)		2.99 (1.01)
dresses well		3.59 (0.91)		3.44 (0.91)
healthy		3.82 (0.81)		3.81 (0.88)
FO	0.86		0.90	
wanting to have children		3.36 (1.31)		3.12 (1.52)
being a good mother/father		3.42 (1.13)		3.67 (1.30)
likes children		3.71 (1.07)		3.72 (1.21)
family-oriented		3.39 (0.99)		3.61 (1.20)

Item of each dimension	a_{T1}	M_{T1} (SD_{T1})	a_{T2}	M_{T2} (SD_{T2})
IQ	0.72		0.78	
intelligent		4.34 (0.68)		4.29 (0.67)
educated		4.19 (0.69)		4.15 (0.79)
sharp		3.56 (0.87)		3.53 (0.92)
clever		3.42 (0.96)		3.45 (0.93)
eloquent		3.62 (0.87)		3.69 (0.90)
interesting		4.19 (0.69)		4.03 (0.76)
C	0.63		0.74	
creative		3.55 (0.87)		3.52 (0.86)
broad-minded		3.21 (0.93)		3.13 (0.93)
inventive		3.94 (0.67)		3.65 (0.86)
original		3.38 (0.84)		3.42 (0.90)
H	0.72		0.77	
fun		3.74 (0.86)		3.76 (0.94)
good fun		3.83 (0.80)		3.79 (0.92)
good sense of humour		4.43 (0.68)		4.16 (0.81)
shrewd		3.11 (1.02)		3.45 (0.92)
smart		1.77 (0.88)		2.05 (0.96)
outgoing		3.57 (0.70)		3.48 (0.81)
AC	0.73		0.77	
adventurous		3.27 (0.93)		3.35 (0.96)
venturesome		2.74 (0.90)		2.70 (0.97)
masculine		2.52 (1.28)		2.50 (1.25)
assertive		3.59 (0.76)		3.24 (0.86)
self-aware		3.83 (0.70)		3.69 (0.80)
ambitious		3.55 (0.75)		3.60 (0.80)
energetic		2.83 (0.90)		2.83 (0.94)
confident		3.87 (0.70)		3.74 (0.75)
dominance		2.22 (0.93)		2.27 (0.95)

Note. WT = warmth-trustworthiness, VA = vitality-attractiveness, SR = status-resources, FO = family-orientation, IQ = intelligence, C = creativity, H = humour and AC = adventurousness-confidence.

Perceived Change of Ideal Partner Preferences

In order to assess how participants perceived to have changed in their ideals, we presented participants the date of their initial participation (MM/YYYY). We asked them to

think back to this former stage of life and to remember exactly what was important to them in an ideal partner at that time. We then asked participants to compare former thoughts and attitudes towards an ideal partner with their current perspective. We presented eight different preference dimensions (corresponding to the factors from the pre-test) which were illustrated with four attributes each (Table S4). Participants indicated whether they perceived to have changed on a 5-point Likert-scale ranging from -2 (*a lot less important than earlier*) to +2 (*a lot more important than earlier*).

Results

Since we decided to only interpret results based on the re-assessment of BSDS participants, we were not able to exactly follow our preregistered analyses. Therefore, only our design and hypotheses can be regarded as preregistered (https://osf.io/x7rma/?view_only=cc43884bd1744047a8afb93106c42c0e). Our analyses were run using R 4.0.2 (R Core Team, 2020b), using the packages psych 2.0.12 (Revelle, 2020), multicon 1.6 (Sherman, 2015), and lmerTest 3.1-2 (Alexandra Kuznetsova et al., 2017), and sjPlots 2.8.7 (Daniel Lüdecke, 2021). Study materials, data and code can be found on the OSF (https://osf.io/z6yaj/?view_only=9680e71797134a9db683a6001494f8fe).

Rank-Order Stability

To estimate the stability of ideal partner preferences across 13 years, we correlated participants' initial (T1) with their current preferences (T2) using a Pearson product-moment correlation, separately for the eight preference dimensions (Table 3). Coefficients ranged from $r = .31$ (for WT) to $r = .47$ (for SR and AC), with a mean correlation of $r = .42$. Men's ($r = .44$) and women's ($r = .36$) mean retest correlation was comparable ($p = .502$, Table S5). These results suggest substantial positive correlations between T1 and T2 preferences, supporting H1.

Table 3.

Means, Standard Deviations, Effect Sizes, and Correlations of T1 and T2 Importance Ratings

	M_{T1} (SD_{T1})	M_{T2} (SD_{T2})	p	d	r [95% CI]
WT	4.16 (.45)	4.25 (.59)	.035	.15	.31 [.19; .43]
VA	3.64 (.53)	3.55 (.59)	.030	.15	.44 [.32; .55]
SR	2.74 (.61)	2.88 (.62)	.002	.22	.47 [.36; .57]
FO	3.47 (.95)	3.53 (1.16)	.452	.05	.46 [.34; .56]
IQ	3.89 (.52)	3.86 (.58)	.521	.05	.37 [.25; .49]
C	3.52 (.57)	3.43 (.67)	.071	.13	.40 [.27 .51]
H	3.41 (.53)	3.45 (.61)	.366	.06	.40 [.27; .51]
AC	3.16 (.50)	3.10 (.55)	.143	.10	.47 [.35; .57]

Note. p = p-values of two-sided t-tests in which we compared participant's mean preferences at T1 and T2. WT = warmth-trustworthiness, VA = vitality-attractiveness, SR = status-resources, FO = family-orientation, IQ = intelligence, C = creativity, H = humour, AC = adventurousness-confidence.

Profile Correlations

For further exploration, we calculated the profile correlation between T1 and T2 ideals dimension-wise. The overall correlation of r = .73 (p <.001) reveals that the pattern of which traits were considered more or less important showed high temporal consistency. Because a substantial part of this association could be due to a normative component of preferences, we also investigated the distinctive profile correlation. Here, an average profile of T1 and T2 preferences is calculated, regressed on each individual profile and the residuals of these regressions are then correlated, resulting in the distinctive profile correlation. The distinctive stability was somewhat smaller, yet still considerable in magnitude (r = .53, p <.001). Additionally, we calculated the profile correlation between T1 and T2 preferences on an item level (overall correlation: r = .62, p <.001; distinctive stability: r = .40, p < .001). Results were comparable for men and women (see S2B). These results corroborate our finding that ideals possess considerable stability from T1 to T2.

Mean-Level Changes of Ideal Partner Preferences

We explored mean-level changes for each dimension using paired sample t-tests (two-sided). We found significant increases in participants' preferences for the dimensions WT and SR, and decreases for VA over time. We found no other significant mean-level

changes (Table 3). These results support H2 in that participants' preference for SR increased over time, but not H3, as there was no increase for FO.

Exploration of sex differences (S2B) revealed that overall, women reported significantly higher ideals ($b = .12$, 95% CI [.02; .21], $p = .017$). On the specific dimensions, women reported a higher preference than men for SR ($b = .36$, 95% CI [.20; .53], $p <.001$), AC ($b = .32$, 95% CI [0.18; .46], $p <.001$), and IQ ($b = .16$, 95% CI [.01; .32], $p = .04$), and a lower on the dimension VA ($b = -.29$, 95% CI [-.43; -.14], $p <.001$).

Age Effects

We investigated the effect of age on mean-level changes across all dimensions using multilevel models. We predicted participants' preferences with the time point ($0 = T1$, $1 = T2$), age (z-standardised) and their interaction, while including a random effect for participants because of the repeated measurement. For SR, we found a significant main effect of time point and interaction of participants' age and time point (Table 4). In other words, the previously described increase over time in SR preferences replicated, yet, as predicted in H2.1, the significant interaction suggested a stronger increase over time for younger participants (Figure 1C, red vs. blue line).

For the dimension FO, we found a significant effect of age and its interaction with time point. Put differently: Contradicting H3, we did not find an overall increase in preference for FO, but as predicted in H3.1, preferences for FO increased from T1 to T2 for younger participants. Interestingly, younger compared to older participants already had a higher preference at T1 (Figure 1B).

Exploratorily, we investigated the effect of age on all other dimensions. Preferences for VA decreased, especially for older participants (Figure 1D). Further, significant main effects of age for IQ and H suggested decreased preferences for both dimensions when being older (Figure 1E, F).

Table 4

Multilevel models investigating Age Effects on each Dimension

	WT		VA		SR		FO	
Predictors	b [95% CI]	p	b [95% CI]	p	b [95% CI]	p	b [95% CI]	p
time point	0.09 [0.01 – 0.18]	0.034	-0.09 [-0.17 – -0.01]	0.027	0.14 [0.05 – 0.22]	0.002	0.06 [-0.09 – 0.21]	0.442
age at T1	-0.02 [-0.10 – 0.05]	0.527	0.06 [-0.02 – 0.14]	0.135	0.08 [-0.01 – 0.16]	0.078	-0.27 [-0.41 – -0.14]	<0.001
time point *age	-0.05 [-0.14 – 0.03]	0.243	-0.11 [-0.19 – -0.03]	0.007	-0.10 [-0.18 – -0.01]	0.029	-0.24 [-0.39 – -0.09]	0.002

	IQ		C		H		AC	
Predictors	b [95% CI]	p	b [95% CI]	p	b [95% CI]	p	b [95% CI]	p
time point	-0.03 [-0.11 – 0.06]	0.520	-0.09 [-0.18 – 0.01]	0.070	0.04 [-0.05 – 0.13]	0.365	-0.06 [-0.13 – 0.02]	0.140
age at T1	-0.09 [-0.17 – -0.02]	0.016	0.03 [-0.06 – 0.11]	0.542	-0.08 [-0.16 – -0.00]	0.040	-0.02 [-0.09 – 0.06]	0.645
time point *age	-0.03 [-0.12 – 0.05]	0.436	-0.02 [-0.12 – 0.07]	0.604	-0.04 [-0.13 – 0.05]	0.356	-0.06 [-0.13 – 0.01]	0.113

Note. WT = warmth-trustworthiness, VA = vitality-attractiveness, SR = status-resources, FO = family-orientation, IQ = intelligence, C = creativity, H = humour, AC = adventurousness-confidence. Full models can be found in our supplement (S2C).

Figure 1

Participant's Mean Ideal Partner Preferences at T1 and T2 Divided into Different Age Groups
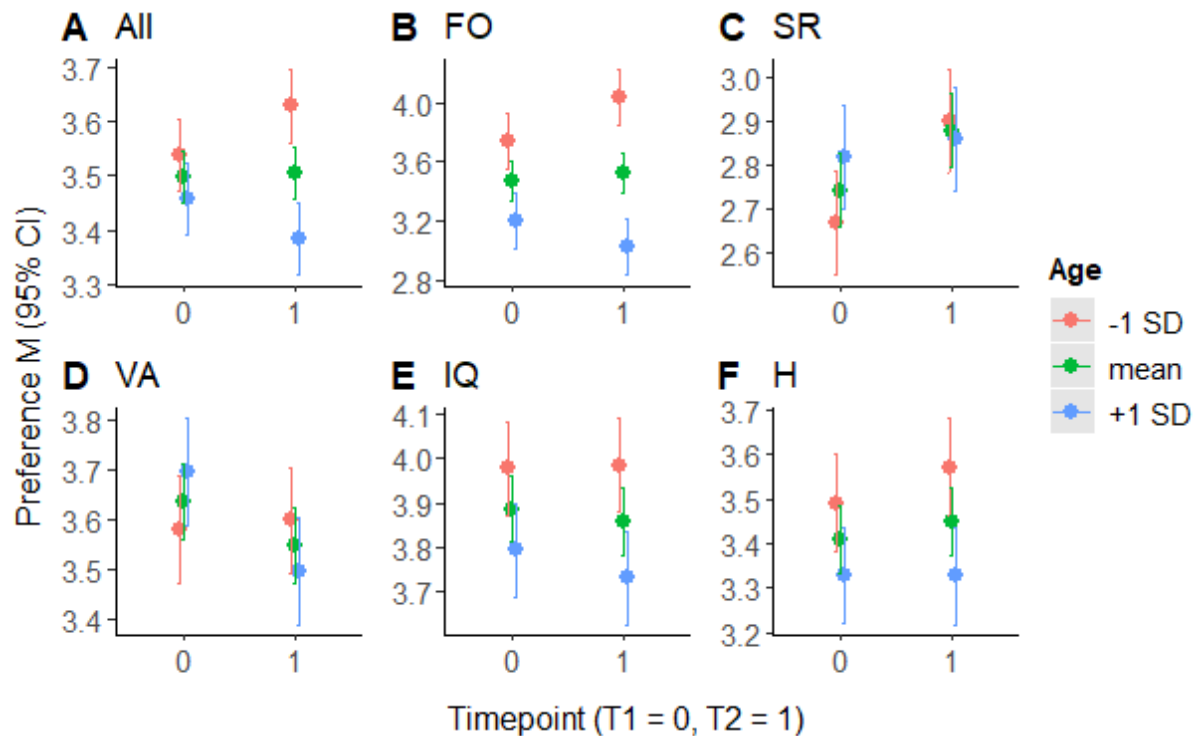
Note. On the y-axis, participants mean preference (95% CI) across all dimensions (A), and separately for the dimension FO (B), SR (C), VA (D), IQ (E) and H (F) plotted separately for each time point (0 = T1, 1 = T2) on the x-axis. To facilitate understanding of our plot, participants were divided into three age groups for illustrating purposes (red = *Mean* age -1 *SD*, green = *Mean* age, blue = *Mean* age +1 *SD*). However, we analysed age continuously. Note that the y-axis differs between each plot.

Participants with vs. without Children

We investigated the effect of parenthood on participants' preferences. After excluding participants who already had children at T1 ($n = 34$, 16%), we divided our sample into two groups: $n = 63$ participants (37%) without children or the intention to have a child and $n = 107$ participants (63%) who have had at least one child since T1 or the intention to have a child (Table S2), henceforth referred to as participants with vs. without children.

We correlated T1 and T2 preferences dimension-wise separately for each group and compared correlation coefficients using a two-sided *z*-test. Overall, mean rank-order stability

was lower in participants with, as compared to without children (r_w = .32 vs. r_{wo} = .54, p = .036). Table 5 depicts rank-order stabilities for each dimension. As predicted in H4, rank-order stability for SR was lower in participants with, as compared to without children. Their rank-order stability was also lower for WT, IQ, C and H, but not in the remaining three dimensions.

Table 5

Means, Standard Deviations, Effect Sizes and Correlations of T1 and T2 Importance Ratings Separately for Participants with and without Children

	with children (n =107)					without children (n = 63)					
	M_{T1w} (SD_{T1w})	M_{T2w} (SD_{T1w})	p_w	d_w	r_w	M_{T1wo} (SD_{T1wo})	M_{T2wo} (SD_{T1wo})	p_{wo}	d_{wo}	r_{wo}	p_c
WT	4.19 (0.43)	4.35 (0.58)	.011	.25	.25	4.08 (0.53)	4.19 (0.52)	.098	.21	.57	.001
VA	3.64 (0.55)	3.55 (0.61)	.106	.16	.47	3.61 (0.53)	3.59 (0.52)	.703	.05	.57	.190
SR	2.68 (0.64)	2.87 (0.62)	.007	.27	.37	2.74 (0.58)	2.83 (0.63)	.185	.17	.60	.006
FO	3.85 (0.72)	4.23 (0.74)	<.001	.40	.14	2.89 (1.05)	2.58 (1.09)	.055	.25	.30	.285
IQ	3.94 (0.51)	3.90 (0.58)	.556	.06	.28	3.85 (0.56)	3.88 (0.55)	.686	.05	.57	.005
C	3.50 (0.57)	3.42 (0.68)	.221	.12	.34	3.5 (0.59)	3.48 (0.59)	.726	.04	.59	.006
H	3.53 (0.48)	3.53 (0.59)	.960	.00	.28	3.28 (0.53)	3.41 (0.62)	.080	.22	.52	.023
AC	3.18 (0.51)	3.13 (0.56)	.367	.09	.43	3.12 (0.53)	3.08 (0.58)	.553	.08	.54	.250

Note. The lower case "w" refers to the group of participants with children, the lower case "wo" refers to participants without children. To show how preferences developed for each group, mean level changes are displayed. However, only results of multilevel models are interpreted for these exploratory analyses. *z*-transformed correlation coefficients of each group are compared using a z-test with the column p_c referring to the *p*-values of each comparison. WT = warmth-trustworthiness, VA = vitality-attractiveness, SR = status-resources, FO = family-orientation, IQ = intelligence, C = creativity, H = humour, AC = adventurousness-confidence.

Exploratorily, we investigated mean-level changes dimension-wise and across all items by running multilevel models. We predicted ideals with the time point (0 = T1, 1 = T2), whether participants have children (0 = without children, 1 = with children), their interaction, and a random effect taking the repeated measurement into account (S2E). Overall, participants with children reported higher ideals (b = .18, 95% CI [.07; .29], p<.001). Dimension-wise investigation revealed that participants with children had a significantly higher preference for the dimension H (b = .25, 95% CI [.08; .43], p = .004). We did not find further significant effects except for FO (time point: b = -.31, 95 % CI [-.58; -.05] p = .023; having a child: b = .96, 95% CI [.69; 1.23], p <.001; their interaction: b = .70, 95% CI [.36; 1.03], p <.001), suggesting that the importance of FO increases over time, but only for participants with children. Participants without children placed less importance on this dimension at both assessments (Figure 2).

Figure 2

FO Preference at T1 and T2 for Participants with vs. without Children
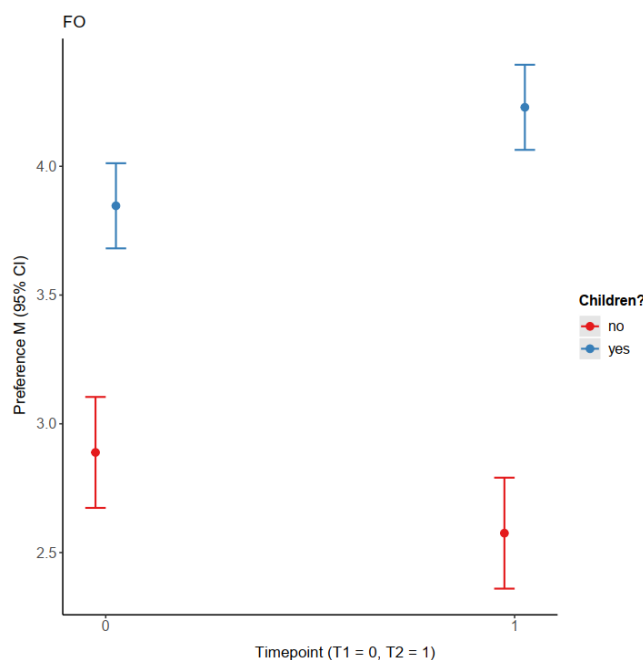
Note. The dots depict participants' mean preference for FO at each time point (0 = T1, 1 = T2). The lines crossing each dot display the 95% CI. In red, participants without children (n = 63, 37%) and in blue participants who have had at least one child since T1 (n = 107, 63%). 34 (16%) participants who already had a child at T1 were not included in these analyses. The figure shows that participants' preference for FO increases from T1 to T2, but only in the group of participants with children. Participants without children place less importance on FO which does not change significantly over time.

Insight into Changes of Ideal Partner Preferences

Finally, we investigated whether participants' have insight into how their preferences have changed. First, we subtracted participants' T1 from their T2 preferences dimension-wise, indicating the actual change of preferences. Descriptively, participants perceived themselves to place less importance on the dimensions VA, SR, C, and AC, and increased importance on the dimensions WT, FO, IQ and H over time. Importantly, for all dimensions except FO, about half of the sample did not perceive themselves to have changed (Table 6). Second, to estimate whether participants have insight into their actual changes, we correlated their perceived change with their actual change dimension-wise. We found a mean accuracy correlation of $r = .20$. Across dimensions, coefficients ranged from $r = .09$ to $r = .45$, suggesting considerable variation in how much insight people had into how much their preferences had changed since T1. For example, as shown in Table 6, participants were quite accurate regarding their change in preference for FO, but had almost no insight into their changes in preference for IQ. In contrast to H5.1, participants did not show any insight regarding their preference change for SR: Preferences increased over time, but participants perceived to have ascribed less importance to this dimension. In line with H5.2, participants accurately perceived a decreased preference for VA over time. We found no intra-individual differences in participants' insight into how their preferences have changed when investigating effects of age or sex and their respective interaction (S2F).

Table 6

Participants' Actual and Perceived Change of each Preference Dimension and their Correlations

	actual change	perceived change	responses of perceived change (%)					Insight	
	M (*SD*)	*M* (*SD*)	-2	-1	0	1	2	*r* [95% CI]	*P*
WT	0.09 (0.62)	0.43 (0.72)	0.0	3.4	59.8	27.0	9.8	.20 [0.06; 0.33]	.005
VA	-0.09 (0.59)	-0.27 (0.72)	2.9	32.8	54.4	8.3	1.5	.21 [0.07; 0.33]	.003
SR	0.14 (0.63)	-0.11 (0.82)	4.9	22.6	52.9	17.7	2.0	.09 [-0.04; 0.23]	.178
FO	0.06 (1.11)	0.08 (1.20)	12.8	15.2	37.3	20.6	14.2	.45 [0.34; 0.56]	<.001
IQ	-0.03 (0.62)	0.21 (0.62)	1.0	3.4	73.5	17.7	4.4	.09 [-0.05; 0.22]	.208
C	-0.09 (0.69)	-0.12 (0.75)	2.5	25.5	54.9	15.7	1.5	.24 [0.10; 0.36]	.001
H	0.04 (0.63)	0.05 (0.63)	1.0	12.3	69.6	15.2	2.0	.14 [0.01; 0.28]	.039
AC	-0.06 (0.54)	-0.16 (0.80)	1.5	33.3	48.5	13.2	3.4	.18 [0.05; 0.31]	.009

Note. Negative values indicate a decreased importance and positive values indicate an increased importance of each corresponding preference dimension. WT = warmth-trustworthiness, VA = vitality-attractiveness, SR = status-resources, FO = family-orientation, IQ = intelligence, C = creativity, H = humour, AC = adventurousness-confidence. *r* [95% CI] = correlation coefficient and 95% confidence interval between actual and perceived changes. *p* = *p*-values of correlation coefficients.

Discussion

In this study, employing unique longitudinal data across 13 years, we investigated stability (i.e. retest and profile correlations) and change (i.e. mean-level changes) and whether individuals possess insight into how their preferences have changed (i.e. correlations of perceived changes with actual changes).

Stability and Change of Ideal Partner Preferences

Regarding our first hypothesis (H1), our results suggested a considerable overall stability of participants' ideals of *r* = .42, corresponding to a medium-sized to large effect (Cohen, 2013; Gignac & Szodorai, 2016). This stability is smaller than coefficients attained after 5 months (Gerlach, Arslan, et al., 2019), yet roughly comparable to coefficients found after 3 years (Bleske-Rechek & Ryan, 2015). These results are in a range that has previously been reported for the rank-order stability of personality traits (around *r* = .60 for a retest interval of 6.7 years, Roberts & DelVecchio, 2000; *r* = .33 for an interval of 11 years in a more diverse sample, Atherton et al., 2021). Further, investigating participants' profiles

revealed that patterns of which traits individuals preferred more or less were surprisingly stable, with overall profile correlations exceeding $r = .70$. These profile stabilities were only slightly reduced when accounting for normative components (e.g. most people value WT more than SR) by employing distinct profile correlations, suggesting that idiosyncratic patterns in what people value in a romantic partner show high temporal consistency.

We then examined effects of parenthood on the stability of preferences. As put forward in hypothesis four (H4), we found that the stability of preferences for SR was lower in participants who became parents over the 13 year study period or who had intentions to become a parent at the time of the re-assessment, compared to participants without (the intention to have) children. We assumed that these shifts in partner preferences could be related to shifting priorities and efforts according to different life stages (cf. Del Giudice et al., 2016; Heckhausen et al., 2010, 2019), with parenthood potentially being of particular importance. As having a partner who is able to provide resources facilitates founding a family and raising children, (the decision to) becoming a parent may alter one's preference for SR, explaining the lower stability. Yet, parenthood was also related to the stability of other preference dimensions, suggesting that the decision to become a parent has the potential to shake up how we picture our ideal partner more generally.

We also investigated mean-level changes. In line with our second hypothesis (H2), participants placed higher importance on SR over time and this increasing preference was stronger for younger participants. Further, although effects were small (Cohen's $d < .20$), participants placed more importance on WT and less on VA over time. Our third prediction (H3), an increase in FO, was only partly supported: Over time, the preference for FO only increased for younger individuals, yet compared to older participants, younger individuals already reported a higher preference for FO at the initial assessment. Further exploration revealed that participants without children generally placed less importance on FO, whereas the preference for FO increased over time for those with children. While this might be a mere cohort effect, this finding could also be interpreted in light of age-graded opportunity structures and/or developmental deadlines (Wrosch & Heckhausen, 2005). For example, younger participants might picture themselves as likely to found a family in the future, whereas older participants had already begun to ponder a possible life without children, either because they already considered themselves to be beyond the ideal age for having children or were pessimistic about finding a suitable partner for such an endeavour.

Insight into Preference Change

Over the 13 years study period, preferences for SR and WT increased, and decreased for VA – but were these changes also mirrored in participants' perceptions? Descriptively, participants perceived to have increased their preference for WT and decreased for VA and SR. They also perceived increases in FO, IQ and H and decreases in AC and C. Interestingly, around 50% of participants did not perceive to have changed their ideals in any dimension, except for FO, where only 37% of participants believed that their preferences had not changed. These patterns dovetail with results by Sprecher and colleagues (2018): Around half of their sample perceived not to have changed their ideals, except for "good parenting potential", a variable close to FO.

When investigating whether perceptions correspond to actual changes, we found a small positive effect. Yet, insight varied considerably between the different dimensions: The most insight was found for FO and the least for SR and IQ. Contradicting our fifth hypothesis (insight into changes for SR, H5.1), participants believed to have decreased in their preference, when in fact they increased over time. Yet, in line with the second part of this prediction (H5.2), participants showed some insight into changes in their preference for VA, even though the perception of change appears stronger than the actual change. Interestingly, age and sex were not related to participants' insight.

The present results for perceptions are in line with previous research (Bleske-Rechek et al., 2009) that found participants to predict that they would value intrinsic characteristics (i.e. WT, FO) more and appearance (i.e. VA) less over time. At the same time, perceptions of change were somewhat exaggerated and for the most part only achieved modest accuracy (FO being a notable exception), showing that perceptions do not necessarily correspond to actual changes. These results highlight the necessity to conduct longitudinal studies and underscore that intra-individual processes should not be investigated in cross-sectional data: Self-perceptions of change do not reflect actual changes accurately enough to allow them to be used as a substitute.

Strengths

The longitudinal design of this study, covering 13 years, makes it unique among studies on the stability and change of partner preferences, which have so far investigated much shorter time periods. Even over this long timespan, our sample size remained fairly

large. A special feature of this sample is that it is a community sample not restricted to the typical student population. Finally, we used comprehensive measures of participants' ideals at both assessments and complementary indices to investigate their stability and change.

Limitations and Future Directions

Even though our community sample was arguably more diverse than the typical student sample, it was still highly educated and came from a Western background. Future studies should strive to recruit participants with more diverse educational backgrounds, ideally also from non-Western countries (Henrich et al., 2010a). Further, although the large retest period is unique and allowed exploring possible effects of life-events, the fact that we only had two assessments available precludes an in-depth analysis of further factors that might have driven preference change. Future research should include multiple assessments of preferences and important events (e.g. parenthood; entering, Gerlach, Arslan, et al., 2019 or ending relationships; experiences of romantic rejection and acceptance, Charlot et al., 2020). Finally, we were not able to fully stick to our preregistered analyses, which is why only our hypotheses and methods can be regarded as preregistered. In particular, the diverging assessment of initial ideals between the two samples led to larger problems than anticipated, which led us to the decision to analyse both samples separately and interpret results based on the BSDS only.

Conclusion

We provide evidence that ideal partner preferences are considerably stable over 13 years, with some changes being related to life events such as parenthood. The importance of a partner with status and resources increased over time, with this increase being stronger for younger individuals. For some preferences, participants knew how they had changed over time. Future research should investigate further factors influencing stability and change of ideals as well as the factors facilitating or hindering insight into such changes.

References

Alexander, R. D. (1987). *The biology of moral systems*. Aldine de Gruyter.

Anusic, I., & Schimmack, U. (2016). Stability and change of personality traits, self-esteem, and well-being: Introducing the meta-analytic stability and change model of retest correlations. *Journal of Personality and Social Psychology*, *110*(5), 766–781. https://doi.org/10.1037/pspp0000066

Arslan, R. C., Walther, M. P., & Tata, C. S. (2020). formr: A study framework allowing for automated feedback generation and complex longitudinal experience-sampling studies using R. *Behavior Research Methods*, *52*(1), 376–387. https://doi.org/10.3758/s13428-019-01236-y

Asendorpf, J. B., & Neyer, F. J. (2012). Persönlichkeitsentwicklung. In J. B. Asendorpf & F. J. Neyer (Eds.), *Psychologie der Persönlichkeit* (pp. 263–331). Springer.

Asendorpf, J. B., Penke, L., & Back, M. D. (2011). From Dating to Mating and Relating: Predictors of Initial and Long–Term Outcomes of Speed–Dating in a Community Sample. *European Journal of Personality*, *25*(1), 16–30. https://doi.org/10.1002/per.768

Atherton, O. E., Sutin, A. R., Terracciano, A., & Robins, R. W. (2021). Stability and change in the Big Five personality traits: Findings from a longitudinal study of Mexican-origin adults. *Journal of Personality and Social Psychology.* Advance online publication. https://doi.org/10.1037/pspp0000385

Bleske-Rechek, A., & Ryan, D. E. (2015). Continuity and change in emerging adults' mate preferences and mating orientations. *Personality and Individual Differences*, *72*, 90–95. https://doi.org/10.1016/j.paid.2014.08.033

Bleske-Rechek, A., VandenHeuvel, B., & Vander Wyst, M. (2009). Age Variation in Mating Strategies and Mate Preferences: Beliefs versus Reality. *Evolutionary Psychology*, *7*(2), 147470490900700204. https://doi.org/10.1177/147470490900700204

Bredow, C. A., & Hames, N. (2019). Steadfast Standards or Fluctuating Fancies? Stability and Change in People's Mate Criteria Over 27 Months. *Personality and Social Psychology Bulletin*, *45*(5), 671–687. https://doi.org/10.1177/0146167218794643

Buss, D. M., & Schmitt, D. P. (1993). Sexual Strategies Theory: An evolutionary perspective on human mating. *Psychological Review*, *100*(2), 204–232. https://doi.org/10.1037/0033-295x.100.2.204

Campbell, L., & Fletcher, G. J. O. (2015). Romantic relationships, ideal standards, and mate selection. *Current Opinion in Psychology*, *1*, 97–100. https://doi.org/10.1016/j.copsyc.2015.01.007

Charlot, N. H., Balzarini, R. N., & Campbell, L. J. (2020). The Association Between Romantic Rejection and Change in Ideal Standards, Ideal Flexibility, and Self-Perceived Mate Value. *Social Psychology*, *51*(2), 116–126. https://doi.org/10.1027/1864-9335/a000392

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Costa, P. T., McCrae, R. R., & Löckenhoff, C. E. (2019). Personality Across the Life Span. *Annual Review of Psychology*, *70*(1), 423–448. https://doi.org/10.1146/annurev-psych-010418-103244

Csajbók, Z., & Berkics, M. (2017). Factor, factor, on the whole, who's the best fitting of all? Factors of mate preferences in a large sample. *Personality and Individual Differences*, *114*, 92–102. https://doi.org/10.1016/j.paid.2017.03.044

Del Giudice, M., Gangestad, S. W., & Kaplan, H. S. (2016). Life history theory and evolutionary psychology. In *The handbook of evolutionary psychology: Foundations, Vol. 1, 2nd ed* (pp. 88–114). John Wiley & Sons, Inc.

Fletcher, G. J. O., Simpson, J. A., & Thomas, G. (2000). Ideals, perceptions, and evaluations in early relationship development. *Journal of Personality and Social Psychology*, *79*(6), 933–940. https://doi.org/10.1037/0022-3514.79.6.933

Fletcher, G. J. O., Simpson, J. A., Thomas, G., & Giles, L. (1999). Ideals in intimate relationships. *Journal of Personality and Social Psychology*, *76*(1), 72–89. https://doi.org/10.1037/0022-3514.76.1.72

Gerlach, T. M., Arslan, R. C., Schultze, T., Reinhard, S. K., & Penke, L. (2019). Predictive validity and adjustment of ideal partner preferences across the transition into romantic relationships. *Journal of Personality and Social Psychology*, *116*(2), 313–330. https://doi.org/10.1037/pspp0000170

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, *102*, 74–78. https://doi.org/10.1016/j.paid.2016.06.069

Heckhausen, J., Wrosch, C., & Schulz, R. (2010). A motivational theory of life-span development. *Psychological Review*, *117*(1), 32–60. https://doi.org/10.1037/a0017668

Heckhausen, J., Wrosch, C., & Schulz, R. (2019). Agency and Motivation in Adulthood and Old Age. *Annual Review of Psychology*, *70*(1), 191–217. https://doi.org/10.1146/annurev-psych-010418-103043

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2-3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *2017*, *82*(13), 26. https://doi.org/10.18637/jss.v082.i13

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*(6), 703–722. https://doi.org/10.1177/0146167216637842

Lüdecke, D. (2021). *sjPlot: Data Visualization for Statistics in Social Science* (Version 2.8.7) [R package]. https://CRAN.R-project.org/package=sjPlot

Penke, L., & Asendorpf, J. B. (2008). Beyond global sociosexual orientations: A more differentiated look at sociosexuality and its effects on courtship and romantic relationships. *Journal of Personality and Social Psychology*, *95*(5), 1113–1135. https://doi.org/10.1037/0022-3514.95.5.1113

R Core Team. (2020). *R: A language and environment for statistical programming* [Computer software]. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/

Revelle, W. (2020). *psych: Procedures for Personality and Psychological Research* (Version 2.0.8) [R-package]. Evanston, Illinois, USA. https://CRAN.R-project.org/package=psych

Roberts, B. W., & DelVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin*, *126*(1), 3–25. https://doi.org/10.1037/0033-2909.126.1.3

Schwarz, S., & Hassebrauck, M. (2012). Sex and Age Differences in Mate-Selection Preferences. *Human Nature*, *23*(4), 447–466. https://doi.org/10.1007/s12110-012-9152-x

Shackelford, T. K., Schmitt, D. P., & Buss, D. M. (2005a). Mate preferences of married
persons in the newlywed year and three years later. *Cognition and Emotion*, *19*(8), 1262–
1270. https://doi.org/10.1080/02699930500215249

Shackelford, T. K., Schmitt, D. P., & Buss, D. M. (2005b). Universal dimensions of human
mate preferences. *Personality and Individual Differences*, *39*(2), 447–458.
https://doi.org/10.1016/j.paid.2005.01.023

Sherman, R. A. (2015). *multicon: Multivariate Constructs* (Version 1.6) [R package].
https://CRAN.R-project.org/package=multicon

Simpson, J. A., Fletcher, G. J. O., & Campbell, L. (2001). The Structure and Function of
Ideal Standards in Close Relationships. In G. J. O. Fletcher & M. S. Clark (Eds.),
Blackwell Handbook of Social Psychology: Interpersonal Processes (pp. 86–106).
Blackwell Publishers.

Sprecher, S., Econie, A., & Treger, S. (2018). Mate preferences in emerging adulthood and
beyond: Age variations in mate preferences and beliefs about change in mate preferences.
Journal of Social and Personal Relationships, *36*(10), 3139–3158.
https://doi.org/10.1177/0265407518816880

Wrosch, C., & Heckhausen, J. (2005). Being On-Time or Off-Time: Developmental
Deadlines for Regulating One's Own Development. In *Thinking time: A multidisciplinary
perspective on time* (pp. 110–123). Hogrefe & Huber Publishers.

Ideal Partner Preferences are Stable Over 13 Years

Supplement

Julie C. Driebe[1], Julia Stern[1,3], Lars Penke[1,2] & Tanja M. Gerlach[1,2]

[1]Department of Psychology, University of Goettingen

[2]Leibniz ScienceCampus Primate Cognition, Goettingen

[3]Department of Psychology, University of Bremen

S1.

A. Further descriptives

Figure S1
Illustration of our study structure
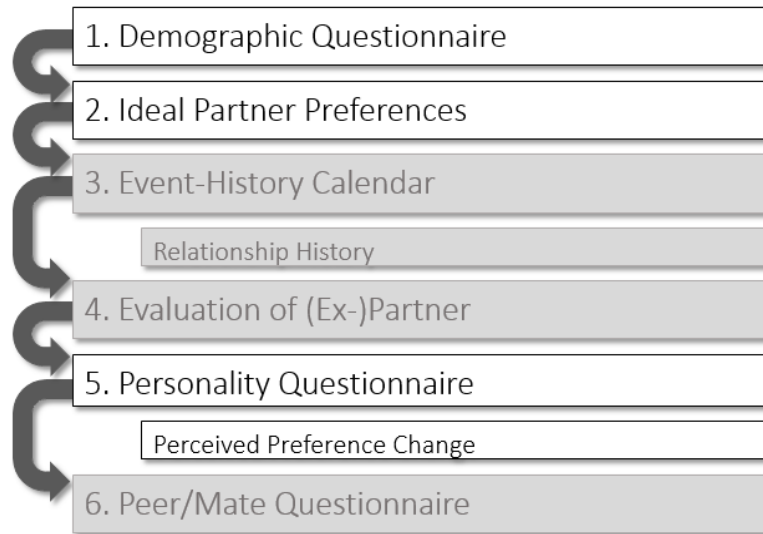
Note. Steps highlighted in light grey are of minor importance for the current study.

Figure S2

Participant flow
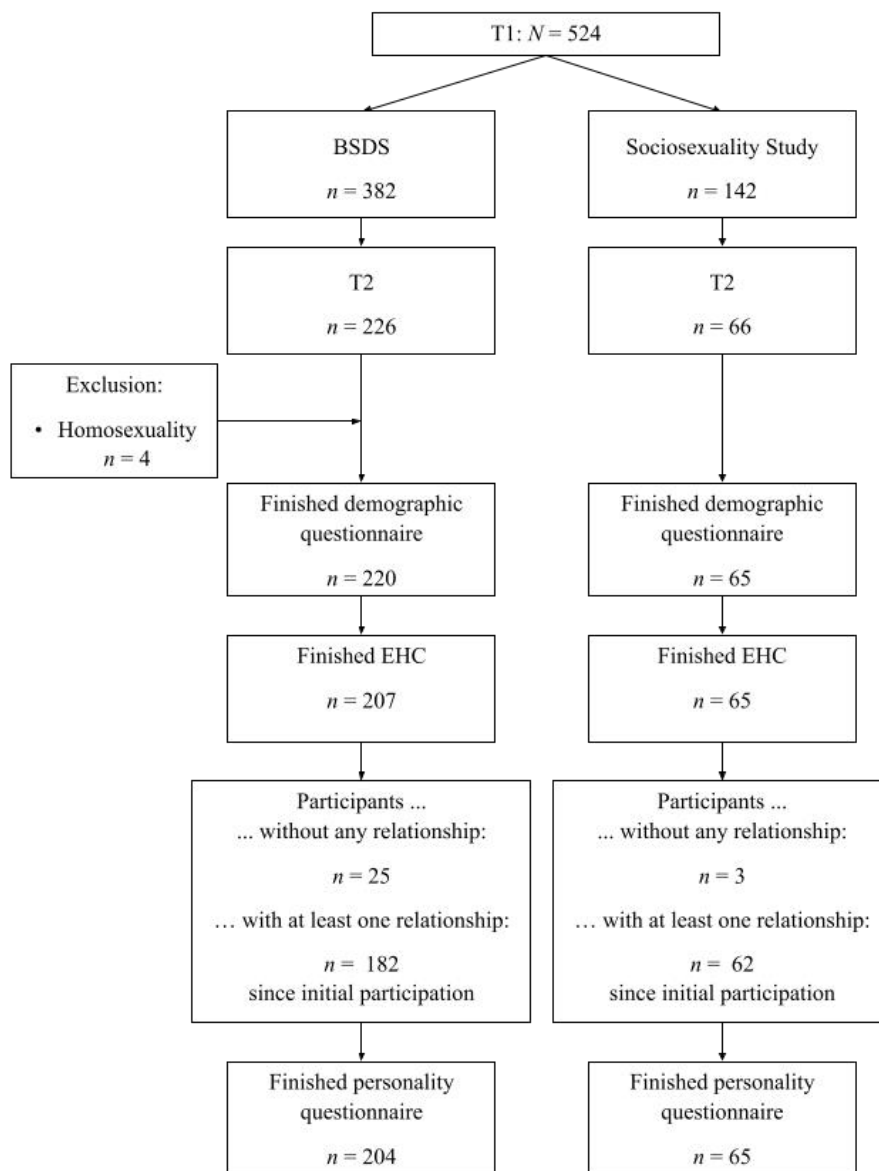
Table S1

Attrition analysis comparing group differences between participants who dropped out from T1 to T2 and between participants who dropped out during T2

Study	T1 values			T2 values		
Participation	Dropouts T1	Re-recruited		Dropouts T2	Final Sample	
n	156	226		22	204	
Variable	*M* (*SD*)	*M* (*SD*)	*p*	*M* (*SD*)	*M* (*SD*)	*p*
Age	33.06 (7.80)	32.56 (7.05)	.53	**41.86 (6.96)**	**46.19 (7.35)**	**.01**
Sex. orientation	1.09 (.29)	1.13 (.36)	.25	1.77 (1.66)	1.08 (.31)	.07
O	3.82 (.53)	3.83 (.50)	.87	3.68 (.89)	3.91 (.58)	.56
C	**3.89 (.60)**	**3.77 (.62)**	**.05**	3.07 (.69)	3.63 (.62)	.14
E	3.57 (.52)	3.47 (.06)	.11	3.65 (.73)	3.42 (.76)	.52
A	3.81 (.50)	3.81 (.54)	.99	3.02 (.99)	3.53 (.60)	.32
N	**2.52 (.70)**	**2.69 (.72)**	**.03**	3.23 (1.18)	2.86 (.85)	.53

Note. P-values are for t-tests (two-sided). Dropouts T1 = group of participants who did not participate at T2; Re-recruited = group of participants who participate at T2; Dropouts T2 = group of participants who had to be excluded or dropped out during T2; Final Sample = final sample at T2; O = Openness; C = Conscientiousness; E = Extraversion; A = Agreeableness; N = Neuroticism; Being assessed using the BFI (Lang, Luedtke, & Asendorpf, 2001) at T2 and the NEO-FFI (Borkenau & Ostendorf, 1993) at T1. Boldface type indicates significant differences with the threshold at the .05.

Table S2

Overview of participants with and without children.

Variable	Sample at T2		Women		Men	
	Yes *n*	No *n*	Yes *n*	No *n*	Yes *n*	No *n*
Child at T1	34 (16%)	170 (83%)	18 (9%)	86 (42%)	16 (8%)	84 (41%)
Child at T2[a]	77 (45%)	93 (55%)	43 (25%)	43 (25%)	34 (20%)	50 (29%)
Pregnant at T2[a]	4 (2%)	166 (98%)	3 (2%)	83 (49%)	1 (1%)	83 (49%)
Currently trying[a]	56 (33%)	114 (67%)	27 (16%)	59 (35%)	29 (17%)	55 (32%)
Child or intention[a]	107 (63%)	63 (37%)	56 (33%)	30 (18%)	51 (30%)	33 (19%)

Note. [a] 34 participants who already had a child at T1 are not included anymore. At T2, participants were additionally asked if they are currently trying to have a/another child. Responses were made on a 7 point-Likert scale ranging from 1 (trying to avoid it) to 7 (trying to). All participants indicating an answer equal or larger than 4 are seen as currently trying to have a/another child.

B. Descriptives of Participants of the Sociosexuality Study

Participants and Procedure

In addition to our main sample, we invited another 142 participants (♀ = 72, ♂ = 70; M = 24.1 years, SD = 2.9, range = 20 - 30) of a former study, the Sociosexuality Study (Penke & Asendorpf, 2008) (T1), to participate in our online study (referred to as T2). T1 took place from 2004 to 2006. We only invited single participants, who made up half of the sample of the Sociosexuality Study, owing to the fact that only their ideal partner preferences were assessed at T1. One of these initial participants also took part in the BSDS, hence we only included this person in the analyses reported in the main text, not in these supplementary analyses. At T2, we were able to re-recruit 66 individuals (dropout rate 54%). Because one person dropped out during our T2 online study, our final sample size consists out of 65 individuals (♀ = 31, ♂ = 34 male; M = 38.7 years, SD = 3.3, range = 34 - 45), with 28% being single at T2, 71% being in a relationship and 2% in an undefined relationship. At T2, 80% of participants had at least a university degree. At T1, none had children. Whereas at T2, 52% had at least one child or were currently pregnant, and 23% were currently trying to have a child. Attrition analyses did not reveal significant group differences between those participants who participated versus those who did not participate at T2 (Table S4).

Table S3

Attrition analysis for the re-assessment of the Sociosexuality Study

Participation	Drop-outs	Re-recruited	
n	76	66	
Variable	*M* (*SD*)	*M* (*SD*)	*p*
Age	23.92 (2.76)	24.36 (3.12)	.37
Sex. orientation	1.20 (.40)	1.23 (.46)	.68
O	3.85 (.68)	3.90 (.54)	.65
C	3.46 (.61)	3.38 (.63)	.45
E	3.56 (.79)	3.61 (.64)	.68
A	3.31 (.56)	3.37 (.57)	.52
N	2.81 (.73)	2.81 (.76)	.97

Note.; O = Openness; C = Conscientiousness; E = Extraversion; A = Agreeableness; N = Neuroticism; Being assessed using the BFI (Lang, Luedtke, & Asendorpf, 2001).

The procedure at T2 was exactly the same as reported for the BSDS in the main text. One important difference at T1 in the Sociosexuality Study is the assessment of ideal partner preferences. Former participants of the Sociosexuality study rated 13 items regarding how strongly they would like a certain characteristic to be pronounced in an ideal partner on a scale ranging from 1 = *very little* to 5 = *as strong as possible*. At T2, they now rated the importance of the same 13 items on a scale ranging from 1 = *very unimportant* to 5 = *very important*. Participants also rated 59 additional items, which were initially assessed in the other sample.

C. Preference Dimensions

Because we could not foresee how many people would participate after around 14 years, we decided to recruit former participants of two different samples, despite the challenging fact that the measure of ideal partner preferences differed between the two samples. To overcome the difficulties of different measurements, we ran a pre-test. A new sample (N = 436, ♀ = 315 [72%], ♂ = 121 [28%]; M = 34.83 years, SD = 13.81 years) rated the initial items of both studies in its importance in an ideal partner and how strongly they would like a certain characteristic to be pronounced. We ran several principal component analyses (separately for each rating scale, a combination of both scales and separately for each sex) with oblimin rotation. An eight factor solution provided the best solution with its factors being: Status-Resources (SR), Warmth-Trustworthiness (WT), Vitality-Attractiveness (VA), Intelligence (IQ), Family Orientation (FO), Humour (H), Creativity (C) and Adventurousness-Confidence (AC). The idea of these preference dimensions was to give us the opportunity not to distinguish between the initial studies in our analysis. Instead we wanted to run our analysis for each preference dimension.

Replication of Preference Dimensions

In a confirmatory factor analysis (CFA), we investigated the fit of our assumed factor structure. Goodness of fit indexes revealed a mixed pattern: (χ^2 F(2441) = 6106.77, p > .001; standardized root mean square residuals [SRMR] = .099; root mean square errors of approximation [RMSEA] = .075, 90% CI[.072; .077]; comparative fit index [CFI] =.694; Tucker-Lewis Index [TLI] = .680). Whereas RMSEA and SRMR fell below an acceptable range, TLI and CFI fell below the threshold of .90. Because fit indices only marginally increased in further exploration (i.e. excluding items with factor loadings smaller than .40) and with our initial factor solution being a compromise between different assessment methods making a better fit unlikely, we calculated participant's mean preferences based on our initial factor structure (Table 2).

Table S4.

Exemplary items of each preference dimension. The first item is always an item used in the initial Sociosexuality Study with the rest of the items deriving from the former BSDS. One exception is the dimension humour which does not contain any item of the former Sociosexuality Study.

Dimension	Content
Status-Resources (SR)	1. occupational success/good occupational prospect 2. wealthy 3. financially secure 4. successful
Warmth-Trustworthiness (WT)	1. faithfulness/reliability 2. understanding 3. sensitive 4. trustworthy
Vitality-Attractiveness (VA)	1. physical attractiveness/sex-appeal 2. erotic 3. sexy 4. arousing
Family-Orientation (FO)	1. parental abilities/whish for children 2. likes children 3. family-oriented 4. being a good mother/father
Intelligence (IQ)	1. intelligence 2. educated 3. sharp 4. clever
Humour (H)	1. fun 2. good sense of humour 3. shrewd 4. good fun
Creativity (C)	1. creativity 2. broad-minded 3. inventive 4. original
Adventurousness - Confidence (AC)	1. adventurousness/activity 2. venturesome 3. assertive 4. confident

S2. Robustness Checks

A. Rank-order Stability

Figure S3.

Plots of T1 preferences predicting T2 preferences, separately for each dimension
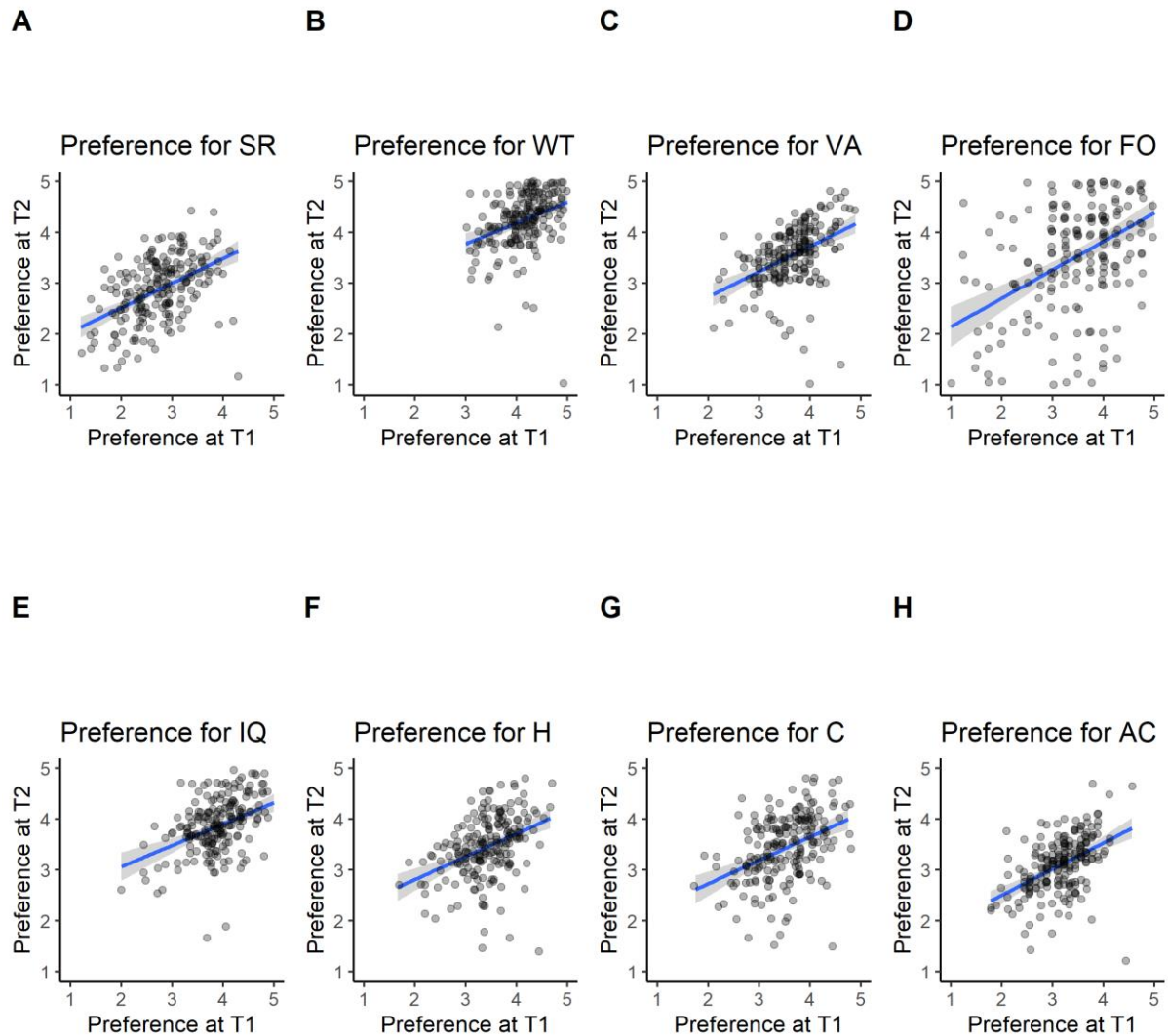
Note. The solid line in blue represents simple slopes derived from the regression models of former participants of the BSDS. The grey area around each line represents the 95% confidence interval.

Table S5

Rank-order stability reported separately for each sex.

	male (*n* = 100)		female (*n* = 104)	
	r [95% CI]	*p*	*r* [95% CI]	*p*
WT	.37 [.18; .52]	<.001	.25 [.06; .42]	.012
VA	.44 [.26; .58]	<.001	.33 [.15; .49]	.001
SR	.55 [.39; .67]	<.001	.33 [.14; .49]	.001
FO	.48 [.31; .61]	<.001	.44 [.28; .59]	<.001
IQ	.43 [.25; .58]	<.001	.31 [.12; .47]	.002
C	.35 [.17; .52]	<.001	.43 [.26; .58]	<.001
H	.40 [.22; .55]	<.001	.39 [.21; .54]	<.001
AC	.48 [.32; .62]	<.001	.39 [.21; .54]	<.001
Overall	.44		.36	.504

Note. WT = warmth-trustworthiness, VA = vitality-attractiveness, SR = status-resources, FO = family-orientation, IQ = intelligence, C = creativity, H = humour, AC = adventurousness-confidence. *p* = *p*-values for each dimension indicate whether correlation coefficients are significant. The *p*-value in the column overall, indicated whether the mean rank-order stability differs between the sexes.

B. Profile Stability (separately for each sex)

As a robustness check, we calculated the profile stability between T1 and T2 preferences separately for each sex. We did so dimension wise (women: overall $r = .70$ (t (104) = 8.54, $p <.001$); distinctive $r = .51$ (t (104) = 8.99, $p <.001$); men: overall $r = .76$ (t (95) = 8.06, $p <.001$); distinctive $r = .50$ (t (95) = 8.56, $p <.001$) and item wise (women: overall $r = .62$ (t (101) = 11.33, $p <.001$); distinctive $r = .35$ (t (101) = 15.86, $p <.001$); men: overall $r = .63$ (t (95) = 10.86, $p <.001$); distinctive $r = .35$ (t (95) = 14.74, $p <.001$)

C. Effect of Sex on Mean-level Changes

Table S6

Multilevel models for the dimension WT, VA, SR and FO

Predictors	WT			VA			SR			FO		
	b	95% CI	p	b	95% CI	p	b	95% CI	p	b	95% CI	p
(Intercept)	4.12	4.02 – 4.23	<0.001	3.78	3.68 – 3.89	<0.001	2.56	2.44 – 2.67	<0.001	3.34	3.13 – 3.55	<0.001
timepoint	0.05	-0.07 – 0.17	0.413	-0.05	-0.16 – 0.07	0.423	0.17	0.04 – 0.29	0.009	0.16	-0.05 – 0.38	0.138
gender	0.07	-0.07 – 0.22	0.336	-0.29	-0.43 – -0.14	<0.001	0.36	0.20 – 0.53	<0.001	0.26	-0.03 – 0.55	0.081
time point*gender	0.08	-0.09 – 0.25	0.352	-0.08	-0.25 – 0.08	0.307	-0.05	-0.23 – 0.12	0.548	-0.21	-0.51 – 0.10	0.181
Random Effects												
σ^2	0.20			0.18			0.20			0.61		
τ_{00}	0.08 session			0.11 session			0.15 session			0.50 session		
ICC	0.30			0.39			0.43			0.45		
N	204 session			204 session			204 session			204 session		
Observations	408			408			408			408		
Marginal R^2 / Conditional R^2	0.020 / 0.312			0.093 / 0.446			0.086 / 0.480			0.008 / 0.455		

Note. b = Estimate, 95% CI = 95% confidence interval, p = p-value. Preferences are predicted with the assessment time point (0 = T1, 1 = T2), gender (0 = male, 1 = female) and their interaction. A random effect is specified for each person.

Table S7

Multilevel models for the dimension IQ, C, AC and H

Predictors	IQ			C			AC			H		
	b	95% CI	*p*	*b*	95% CI	*p*	b	95% CI	*p*	*b*	95% CI	*p*
(Intercept)	3.80	3.70 – 3.91	<0.001	3.53	3.41 – 3.65	<0.001	3.00	2.90 – 3.10	<0.001	3.37	3.25 – 3.48	<0.001
time point	0.01	-0.11 – 0.13	0.829	-0.06	-0.20 – 0.07	0.344	-0.01	-0.11 – 0.10	0.919	0.04	-0.08 – 0.17	0.493
gender	0.16	0.01 – 0.31	0.040	-0.03	-0.20 – 0.14	0.751	0.32	0.18 – 0.46	<0.001	0.08	-0.07 – 0.24	0.297
time point*gender	-0.08	-0.25 – 0.09	0.351	-0.04	-0.23 – 0.15	0.654	-0.10	-0.25 – 0.05	0.193	-0.01	-0.18 – 0.17	0.942
Random Effects												
σ^2	0.19			0.24			0.15			0.20		
τ_{00}	0.11 session			0.15 session			0.11 session			0.13 session		
ICC	0.37			0.39			0.43			0.39		
N	204 session			204 session			204 session			204 session		
Observations	408			408			408			408		
Marginal R^2 / Conditional R^2	0.013 / 0.376			0.007 / 0.394			0.071 / 0.472			0.006 / 0.394		

Note. b = Estimate, 95% CI = 95% confidence interval, *p* = *p*-value. Preferences are predicted with the assessment time point (0 = T1, 1 = T2), sex (0 = male, 1 = female) and their interaction. A random effect is specified for each person.

D. Effect of Age on Mean-level Changes

Table S8

Multilevel models for the dimension WT, VA, SR and FO

Predictors	WT			VA			SR			FO		
	b	95% CI	p	b	95% CI	p	b	95% CI	p	b	95% CI	p
(Intercept)	4.16	4.09 – 4.23	<0.001	3.64	3.56 – 3.71	<0.001	2.74	2.66 – 2.83	<0.001	3.47	3.34 – 3.60	<0.001
time point	0.09	0.01 – 0.18	0.034	-0.09	-0.17 – -0.01	0.027	0.14	0.05 – 0.22	0.002	0.06	-0.09 – 0.21	0.442
age at T1	-0.02	-0.10 – 0.05	0.527	0.06	-0.02 – 0.14	0.135	0.08	-0.01 – 0.16	0.078	-0.27	-0.41 – -0.14	<0.001
time point*age	-0.05	-0.14 – 0.03	0.243	-0.11	-0.19 – -0.03	0.007	-0.10	-0.18 – -0.01	0.029	-0.24	-0.39 – -0.09	0.002
Random Effects												
σ^2	0.19			0.17			0.20			0.59		
τ_{00}	0.08 session			0.14 session			0.18 session			0.36 session		
ICC	0.30			0.45			0.48			0.38		
N	204 session			204 session			204 session			204 session		
Observations	408			408			408			408		
Marginal R^2 / Conditional R^2	0.018 / 0.313			0.016 / 0.462			0.020 / 0.491			0.150 / 0.475		

Note. b = Estimate, 95% CI = 95% confidence interval, *p* = *p*-value. Preferences are predicted with the assessment time point (0 = T1, 1 = T2), age (z-standardised) and their interaction. A random effect is specified for each person.

Table S9

Multilevel models for the dimension IQ, C, AC and H

Predictors	IQ			C			H			AC		
	b	95% CI	p	b	95% CI	p	b	95% CI	p	b	95% CI	p
(Intercept)	3.89	3.81 – 3.96	<0.001	3.52	3.43 – 3.60	<0.001	3.41	3.33 – 3.49	<0.001	3.16	3.09 – 3.23	<0.001
time point	-0.03	-0.11 – 0.06	0.520	-0.09	-0.18 – 0.01	0.070	0.04	-0.05 – 0.13	0.365	-0.06	-0.13 – 0.02	0.140
age at T1	-0.09	-0.17 – -0.02	0.016	0.03	-0.06 – 0.11	0.542	-0.08	-0.16 – -0.00	0.040	-0.02	-0.09 – 0.06	0.645
time point*age	-0.03	-0.12 – 0.05	0.436	-0.02	-0.12 – 0.07	0.604	-0.04	-0.13 – 0.05	0.356	-0.06	-0.13 – 0.01	0.113
Random Effects												
σ^2	0.19			0.24			0.20			0.15		
τ_{00}	0.10 $_{\text{session}}$			0.15 $_{\text{session}}$			0.12 $_{\text{session}}$			0.13 $_{\text{session}}$		
ICC	0.35			0.39			0.38			0.47		
N	204 $_{\text{session}}$			204 $_{\text{session}}$			204 $_{\text{session}}$			204 $_{\text{session}}$		
Observations	408			408			408			408		
Marginal R^2 / Conditional R^2	0.040 / 0.375			0.006 / 0.394			0.034 / 0.396			0.014 / 0.475		

Note. b = Estimate, 95% CI = 95% confidence interval, p = p-value. Preferences are predicted with the assessment time point (0 = T1, 1 = T2), age (z-standardised) and their interaction. A random effect is specified for each person.

E. Effect of Parenthood on Mean-level Changes

Table S10

Multilevel models for the dimension WT, VA, SR and FO

Predictors	WT			VA			SR			FO		
	b	95% CI	p	b	95% CI	p	b	95% CI	p	b	95% CI	p
(Intercept)	4.08	3.96 – 4.21	<0.001	3.61	3.48 – 3.75	<0.001	2.74	2.59 – 2.89	<0.001	2.89	2.67 – 3.10	<0.001
time point	0.10	-0.04 – 0.25	0.161	-0.02	-0.16 – 0.11	0.739	0.09	-0.07 – 0.25	0.263	-0.31	-0.58 – -0.05	0.022
Child at T2	0.10	-0.06 – 0.26	0.201	0.03	-0.15 – 0.20	0.768	-0.06	-0.25 – 0.13	0.537	0.96	0.69 – 1.23	<0.001
time point * Child at T2	0.05	-0.13 – 0.23	0.553	-0.07	-0.25 – 0.10	0.425	0.10	-0.11 – 0.30	0.350	0.70	0.36 – 1.03	<0.001
Random Effects												
σ^2	0.17			0.16			0.21			0.59		
τ_{00}	0.10 $_{session}$			0.16 $_{session}$			0.17 $_{session}$			0.17 $_{session}$		
ICC	0.36			0.50			0.45			0.23		
N	170 $_{session}$			170 $_{session}$			170 $_{session}$			170 $_{session}$		
Observations	340			340			340			340		
Marginal R^2 / Conditional R^2	0.033 / 0.386			0.005 / 0.504			0.016 / 0.458			0.362 / 0.507		

Note. b = Estimate, 95% CI = 95% confidence interval, p = p-value. Preferences are predicted with the assessment time point (0 = T1, 1 = T2) and whether participants have children at T2 (0 = no children, 1 = children) and their interaction. A random effect is specified for each person.

Table S11

Multilevel models for the dimension IQ, C, AC and H

		IQ				C				AC				H	
Predictors	*b*	95% CI	*p*	*b*	95% CI	*p*	*b*	95% CI	*p*	*b*	95% CI	*p*
(Intercept)	3.85	3.72 – 3.99	<0.001	3.50	3.35 – 3.65	<0.001	3.12	2.98 – 3.25	<0.001	3.28	3.14 – 3.42	<0.001
time point	0.03	-0.12 – 0.18	0.730	-0.02	-0.19 – 0.14	0.775	-0.04	-0.18 – 0.10	0.564	0.13	-0.03 – 0.28	0.103
Child at T2	0.09	-0.08 – 0.26	0.308	0.00	-0.19 – 0.20	0.962	0.06	-0.11 – 0.23	0.471	0.25	0.08 – 0.43	0.004
time point * child at T2	-0.06	-0.25 – 0.13	0.509	-0.06	-0.27 – 0.14	0.552	-0.01	-0.18 – 0.16	0.917	-0.13	-0.32 – 0.06	0.185
Random Effects												
σ^2		0.18			0.22			0.16			0.19	
τ_{00}		0.12 session			0.16 session			0.14 session			0.11 session	
ICC		0.39			0.42			0.47			0.37	
N		170 session			170 session			170 session			170 session	
Observations		340			340			340			340	
Marginal R^2 / Conditional R^2		0.003 / 0.390			0.004 / 0.424			0.004 / 0.474			0.031 / 0.392	

Note. b = Estimate, 95% CI = 95% confidence interval, *p* = *p*-value. Preferences are predicted with the assessment time point (0 = T1, 1 = T2) and whether participants have children at T2 (0 = no children, 1 = children) and their interaction. A random effect is specified for each person.

F. Insight into Preference Change

Figure S4

Histogram of participants' perception of change
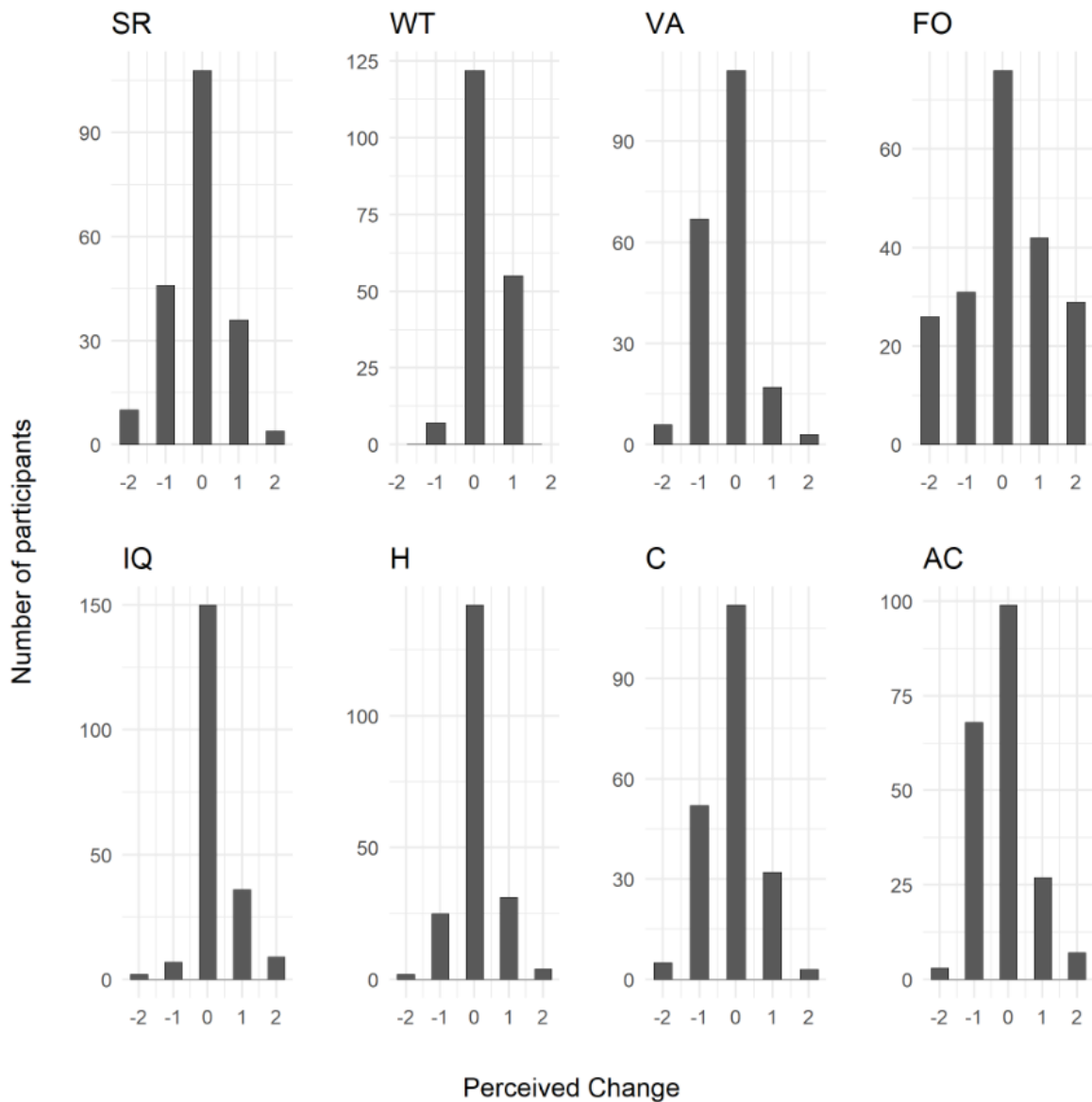
Note. Negative values indicate a reduced importance, zero indicates no change and positive values indicate an increased importance of the corresponding preference dimensions.

Effects of Age

Table S12
Linear models investigating whether age effects participant's insight into preference changes for the dimension WT, VA, SR and FO.

	WT			VA			SR			FO		
Predictors	*b*	95% CI	*p*	*b*	95% CI	*p*	*b*	95% CI	*p*	*b*	95% CI	*p*
(Intercept)	0.09	0.01 – 0.18	0.032	-0.09	-0.17 – -0.01	0.026	0.14	0.05 – 0.23	0.003	0.01	-0.14 – 0.16	0.893
perceived change	0.12	0.03 – 0.21	0.008	0.12	0.04 – 0.20	0.005	0.04	-0.05 – 0.13	0.409	0.50	0.35 – 0.65	<0.001
age	-0.06	-0.14 – 0.03	0.174	-0.11	-0.19 – -0.03	0.006	-0.08	-0.17 – 0.01	0.095	-0.02	-0.17 – 0.14	0.818
perceived change * age	0.01	-0.08 – 0.09	0.884	0.01	-0.08 – 0.09	0.884	0.00	-0.08 – 0.09	0.932	-0.11	-0.26 – 0.05	0.177
Observations	204			204			204			204		
R^2 / R^2 adjusted	0.048 / 0.033			0.079 / 0.065			0.024 / 0.009			0.213 / 0.201		

Note. b = Estimate, 95% CI = 95% confidence interval, *p* = *p*-value. Perceived change and age were *z*-standardised.

Table S13
Linear models investigating whether age effects participant's insight into preference changes for the dimension IQ, C, H and AC.

	IQ			C			H			AC		
Predictors	*b*	95% CI	*p*	*b*	95% CI	*p*	*b*	95% CI	*p*	*b*	95% CI	*p*
(Intercept)	-0.04	-0.13 – 0.05	0.364	-0.09	-0.18 – 0.01	0.070	0.04	-0.05 – 0.13	0.343	-0.06	-0.13 – 0.02	0.127
perceived change	0.05	-0.04 – 0.13	0.292	0.17	0.08 – 0.27	<0.001	0.09	0.00 – 0.18	0.049	0.09	0.02 – 0.17	0.018
age	-0.04	-0.13 – 0.05	0.352	-0.07	-0.16 – 0.02	0.139	-0.04	-0.13 – 0.04	0.319	-0.06	-0.14 – 0.01	0.103
perceived change * age	-0.06	-0.14 – 0.02	0.112	-0.01	-0.09 – 0.07	0.743	0.02	-0.08 – 0.11	0.708	-0.03	-0.11 – 0.05	0.402
Observations	204			204			204			204		
R^2 / R^2 adjusted	0.024 / 0.009			0.067 / 0.053			0.027 / 0.012			0.049 / 0.035		

Note. b = Estimate, 95% CI = 95% confidence interval, *p* = *p*-value. Perceived change and age were *z*-standardised.

Effects of Sex

Table 14
Linear models investigating whether sex effects participant's insight into preference changes for the dimension WT, VA, SR and FO.

Predictors	WT			VA			SR			FO		
	b	95% CI	*p*	*b*	95% CI	*p*	*b*	95% CI	*p*	*b*	95% CI	*p*
(Intercept)	0.05	-0.07 – 0.17	0.440	-0.06	-0.17 – 0.06	0.338	0.16	0.04 – 0.29	0.010	0.17	-0.02 – 0.36	0.082
perceived change	0.13	0.01 – 0.26	0.032	0.08	-0.02 – 0.19	0.129	-0.02	-0.16 – 0.13	0.837	0.65	0.44 – 0.85	<0.001
sex	0.09	-0.08 – 0.26	0.310	-0.06	-0.22 – 0.10	0.469	-0.06	-0.24 – 0.11	0.492	-0.22	-0.49 – 0.05	0.113
perceived change * sex	-0.02	-0.19 – 0.15	0.826	0.08	-0.08 – 0.24	0.324	0.12	-0.06 – 0.31	0.188	-0.26	-0.53 – 0.01	0.061
Observations	204			204			204			204		
R^2 / R^2 adjusted	0.044 / 0.030			0.050 / 0.036			0.020 / 0.005			0.229 / 0.217		

Note. b = Estimate, 95% CI = 95% confidence interval, *p* = *p*-value. Perceived change was *z*-standardised. Sex was dummy-coded (0 = male, 1= female).

Table 15
Linear models investigating whether sex effects participant's insight into preference changes for the dimension IQ, C, H and AC.

Predictors	IQ			C			H			AC		
	b	95% CI	*p*	*b*	95% CI	*p*	*b*	95% CI	*p*	*b*	95% CI	*p*
(Intercept)	0.03	-0.09 – 0.15	0.667	-0.09	-0.23 – 0.04	0.170	0.05	-0.07 – 0.18	0.400	-0.01	-0.11 – 0.10	0.900
perceived change	0.13	0.01 – 0.26	0.033	0.20	0.06 – 0.33	0.005	0.13	0.00 – 0.25	0.045	0.08	-0.02 – 0.19	0.130
sex	-0.09	-0.26 – 0.08	0.283	0.00	-0.18 – 0.19	0.968	-0.02	-0.19 – 0.15	0.815	-0.10	-0.24 – 0.05	0.202
perceived change * sex	-0.14	-0.31 – 0.03	0.098	-0.06	-0.25 – 0.13	0.512	-0.07	-0.24 – 0.11	0.450	0.03	-0.12 – 0.18	0.719
Observations	204			204			204			204		
R^2 / R^2 adjusted	0.027 / 0.012			0.058 / 0.044			0.024 / 0.009			0.041 / 0.027		

Note. b = Estimate, 95% CI = 95% confidence interval, *p* = *p*-value. Perceived change was *z*-standardised. Sex was dummy-coded (0 = male, 1= female).

S3. Results for each study and across studies

A. Rank-Order Stability

In Table S5, we calculated the correlation of T1 and T2 preferences once for both samples (GLIMP), only for the former participants of the Sociosexuality Study. However, results of former participants of the Sociosexuality Study should be interpreted with caution due to the limited number of items assessing T1 preferences (13 items) and the small sample size ($n = 65$).

Table S16

Correlation coefficients between T1 and T2 preferences for the whole sample and separately for each initial study.

	GLIMP		Sociosexuality study	
	r	*95% CI*	*r*	*95% CI*
SR	.39	[.29; .49]	.24	[-.00; .46]
WT	.16	[.04; .27]	.03	[-.22; .27]
VA	.38	[.27; .47]	.19	[-.06; .41]
FO	.40	[.29; .49]	.22	[-.02; .44]
IQ	.22	[.10; .33]	.12	-.12; .36]
H	-	-	-	-
C	.26	[.14; .37]	.02	[-.23; .26]
AC	.33	[.22; .44]	.16	[-.09; .39]
Overall	.32		.14	

B. Profile Correlation

Across both samples, we found a normative profile correlation of $\bar{q} = 0.62$ (t (204) = 19.67, p <.001) and a distinctive profile correlation of $\bar{q} = 0.39$ (t (204) = 23.82, p <.001). For former participants of the Sociosexuality study, we found a normative profile correlation of $\bar{q}$ = 0.17 (t (65) = 2.10, p <.020) and a distinctive profile correlation of $\bar{q}$ = 0.11 (t (65) = 2.34, p <.011).

C. Mean-level Changes

Table S17

Means, standard deviations and effect sizes of T1 and T2 importance ratings reported across studies (GLIMP) and for former participants of the Sociosexuality Study (SOI).

	GLIMP (*n* = 269)				SOI (*n* =65)			
	M_{T1} (*SD*)	M_{T2} (*SD*)	*P*	*d*	M_{T1} (*SD*)	M_{T2} (*SD*)	*p*	*d*
SR	2.86 (.60)	2.92 (.63)	.10	.10	3.21 (.42)	3.00 (.69)	=.02	.29
WT	4.02 (.55)	4.29 (.56)	< .001	.37	3.58 (.61)	4.49 (.73)	<.001	.94
VA	3.64 (.54)	3.59 (.58)	.26	.07	3.64 (.57)	3.80 (.58)	.08	.08
FO	3.40 (.93)	3.56 (1.16)	.03	.14	3.17 (.82)	3.66 (1.27)	=.004	.37
IQ	3.75 (.62)	3.94 (.58)	< .001	.24	3.34 (.71)	4.28 (.70)	<.001	1.01
H	-		-	-	-	3.67 (.54)	-	-
C	3.53 (.63)	3.44 (.65)	.04	.13	3.57 (.79)	3.35 (.93)	.15	.18
AC	3.25 (.56)	3.15 (.55)	.01	.16	3.55 (.64)	3.68 (.85)	.31	.13
Overall	3.46 (.31)	3.52 (.40)	.03	.13				

Note. At T1, importance of humour was only assessed in the BSDS sample but not in the Sociosexuality Study. For the overall T1 and T2 preference, we calculated the mean importance rating of all preference items separately for each timepoint. For the overall correlation we z-transformed each correlation of each preference dimension, calculated the mean z-value and transformed it into a correlation coefficient.

D. Effects of Age

Table S18

Multilevel models investigating whether age effects participant's preference change for the dimension WT, VA, SR and FO reported across both initial samples (GLIMP)

Predictors	WT b	95% CI	p	VA b	95% CI	p	SR b	95% CI	p	FO b	95% CI	p
(Intercept)	4.02	3.96 – 4.09	<0.001	3.64	3.57 – 3.70	<0.001	2.86	2.78 – 2.93	<0.001	3.40	3.28 – 3.52	<0.001
time point	0.27	0.18 – 0.35	<0.001	-0.04	-0.12 – 0.03	0.249	0.07	-0.01 – 0.15	0.095	0.16	0.03 – 0.30	0.015
age at T1	0.11	0.04 – 0.17	0.002	0.04	-0.02 – 0.11	0.218	-0.05	-0.12 – 0.03	0.219	-0.12	-0.23 – 0.00	0.056
time point * age at T1	-0.17	-0.26 – -0.09	<0.001	-0.12	-0.19 – -0.05	0.001	0.03	-0.05 – 0.12	0.402	-0.36	-0.50 – -0.23	<0.001
Random Effects												
σ^2	0.25			0.19			0.23			0.61		
τ_{00}	0.06 $_{session}$			0.12 $_{session}$			0.15 $_{session}$			0.37 $_{session}$		
ICC	0.19			0.39			0.39			0.38		
N	269 $_{session}$			269 $_{session}$			269 $_{session}$			269 $_{session}$		
Observations	538			538			538			538		
Marginal R^2 / Conditional R^2	0.078 / 0.249			0.014 / 0.399			0.006 / 0.395			0.115 / 0.449		

Note. b = Estimate, 95% CI = 95% confidence interval, *p* = *p*-value. Preferences are predicted with the assessment time point (0 = T1, 1 = T2), age at T1 (*z*-standardised) and their interaction. A random effect is specified for each person.

Table S19

Multilevel models investigating whether age effects participant's preference change for the dimension IQ, C, H and AC reported across both initial samples (GLIMP)

		IQ			C			H			AC	
Predictors	b	95% CI	p	b	95% CI	p	b	95% CI	p	b	95% CI	p
(Intercept)	3.75	3.68 – 3.82	<0.001	3.53	3.45 – 3.61	<0.001	3.44	3.36 – 3.52	<0.001	3.25	3.19 – 3.32	<0.001
time point	0.18	0.10 – 0.27	<0.001	-0.10	-0.19 – -0.00	0.040	0.06	-0.02 – 0.15	0.153	-0.10	-0.18 – -0.02	0.011
age at T1	0.05	-0.02 – 0.12	0.187	0.01	-0.06 – 0.09	0.751	-0.09	-0.17 – -0.02	0.019	-0.12	-0.18 – -0.05	<0.001
time point * age at T1	-0.16	-0.25 – -0.07	<0.001	-0.01	-0.10 – 0.08	0.837	-0.06	-0.14 – 0.03	0.188	0.04	-0.04 – 0.12	0.307
Random Effects												
σ^2		0.27			0.31			0.20			0.21	
τ_{00}		0.08 session			0.11 session			0.11 session			0.09 session	
ICC		0.23			0.26			0.37			0.31	
N		269 session			269 session			269 session			269 session	
Observations		538			538			473			538	
Marginal R^2 / Conditional R^2		0.044 / 0.268			0.006 / 0.262			0.056 / 0.404			0.039 / 0.341	

Note. b = Estimate, 95% CI = 95% confidence interval, p = p-value. Preferences are predicted with the assessment time point (0 = T1, 1 = T2), age at T1 (z-standardised) and their interaction. A random effect is specified for each person.

Table S20

Multilevel models investigating whether age effects participant's preference change for the dimension WT, VA, SR and FO reported for former participants of the Sociosexuality Study.

Predictors	WT			VA			SR			FO		
	b	95% CI	p	b	95% CI	p	b	95% CI	p	b	95% CI	p
(Intercept)	3.58	3.42 – 3.75	<0.001	3.64	3.50 – 3.78	<0.001	3.21	3.07 – 3.35	<0.001	3.17	2.91 – 3.43	<0.001
time point	0.91	0.68 – 1.14	<0.001	0.16	-0.02 – 0.34	0.074	-0.21	-0.38 – -0.03	0.020	0.49	0.17 – 0.81	0.003
age at T1	-0.03	-0.19 – 0.14	0.757	-0.01	-0.15 – 0.13	0.936	0.01	-0.13 – 0.15	0.889	0.17	-0.09 – 0.43	0.212
time point * age at T1	-0.12	-0.35 – 0.11	0.295	-0.13	-0.31 – 0.05	0.147	0.83	-0.26 – 0.09	0.358	-0.30	-0.63 – 0.02	0.065
Random Effects												
σ^2	0.44			0.27			0.26			0.88		
τ_{00}	0.01 session			0.06 session			0.07 session			0.26 session		
ICC	0.02			0.19			0.22			0.23		
N	65 session			65 session			65 session			65 session		
Observations	130			130			130			130		
Marginal R^2 / Conditional R^2	0.328 / 0.341			0.046 / 0.230			0.039 / 0.247			0.069 / 0.281		

Note. b = Estimate, 95% CI = 95% confidence interval, p = p-value. Preferences are predicted with the assessment time point (0 = T1, 1 = T2), age at T1 (*z*-standardised) and their interaction. A random effect is specified for each person.

Table S21

Multilevel models investigating whether age effects participant's preference change for the dimension IQ, C and AC reported for former participants of the Sociosexuality Study

Predictors	IQ			C			AC		
	b	95% CI	*p*	*b*	95% CI	*p*	*b*	95% CI	*p*
(Intercept)	3.34	3.17 – 3.51	<0.001	3.57	3.36 – 3.78	<0.001	3.55	3.38 – 3.73	<0.001
time point	0.94	0.71 – 1.17	<0.001	-0.22	-0.51 – 0.08	0.153	0.12	-0.12 – 0.36	0.313
age at T1	-0.05	-0.22 – 0.12	0.549	0.04	-0.17 – 0.25	0.702	-0.18	-0.36 – -0.01	0.044
time point * age at T1	-0.05	-0.28 – 0.18	0.672	-0.00	-0.30 – 0.29	0.976	-0.00	-0.24 – 0.24	0.989
Random Effects									
σ^2	0.44			0.74			0.48		
τ_{00}	0.06 session			0.01 session			0.05 session		
ICC	0.11			0.02			0.10		
N	65 session			65 session			65 session		
Observations	130			130			130		
Marginal R^2 / Conditional R^2	0.315 / 0.392			0.017 / 0.034			0.066 / 0.160		

Note. b = Estimate, 95% CI = 95% confidence interval, *p* = *p*-value. Preferences are predicted with the assessment time point (0 = T1, 1 = T2), age at T1 (*z*-standardised) and their interaction. A random effect is specified for each person.

E. Effects of Parenthood

Rank-Order stability for participants with and without children

We decided not to perform these analyses for former participants of the Sociosexuality Study because of the small sample size. We would have needed to divide the already small sample of 65 participants, in an even much smaller subgroup of participants with vs. without children. These analyses would not provide any reliable insight. Therefore we only report rank-order stabilities for participants with vs. without children across both studies.

Table S22

Means, standard deviations, effect sizes and correlations of T1 and T2 importance ratings separately for participants with and without children

	participants with children at T2 ($n = 155$)					participants without children at T2 ($n = 88$)				
dim	M_{T1} (SD)	M_{T2} (SD)	p	d	r_w	M_{T1} (SD)	M_{T2} (SD)	p	d	r_{wo}
SR	2.83 (.63)	2.93 (.59)	.16	.14	.32	2.87 (.59)	2.87 (.70)	.98	.00	.46
WT	4.00 (.57)	4.37 (.54)	<.001	.48	.09	3.97 (.57)	4.20 (.50)	=.007	.39	.39
VA	3.65 (.55)	3.62 (.58)	.67	.04	.40	3.61 (.54)	3.61 (.53)	.97	.01	.44
FO	3.68 (.74)	4.22 (.72)	<.001	.56	.12	2.86 (1.05)	2.54 (1.10)	.06	.25	.24
IQ	3.75 (.64)	3.99 (.54)	<.001	.30	.10	3.76 (.64)	3.92 (.60)	.10	.25	.43
H	3.53 (.48)	3.53 (.59)	.97	.00	.28	3.28 (.53)	3.41 (.62)	.22	.22	.52
C	3.53 (.64)	3.43 (.64)	.16	.13	.20	3.51 (.65)	3.46 (.61)	.63	.07	.37
AC	3.28 (.57)	3.20 (.54)	.18	.13	.30	3.23 (.61)	3.09 (.60)	.15	.20	.36

Note. Results for the dimension H are only based on participants and items of the former BSDS.

F. Insight into Preference Change

Table S23

Participant's actual and perceived change of each preference dimension as well as the correlation across participants of both former studies.

	actual change		perceived change		responses of perceived change (%)					insight		
	M	*SD*	*M*	*SD*	-2	-1	0	1	2	*r*	*p*	95% CI
SR	0.07	0.68	-0.06	0.80	4.0	20.4	55.0	18.2	2.2	0.15	=.01	[0.03;0.27]
WT	0.27	0.72	0.48	0.73	0.0	3.3	56.5	29.4	10.8	0.20	<.001	[0.08;0.31]
VA	-0.04	0.62	-0.26	0.71	2.6	32.0	55.8	8.2	1.5	0.23	<.001	[0.11;0.34]
FO	0.16	1.17	0.24	120	11.2	11.9	36.1	23.8	17.1	0.50	<.001	[0.40;0.58]
IQ	0.18	0.75	0.24	0.66	1.1	3.7	70.3	19.7	5.2	0.14	=.02	[0.02;0.26]
H	0.04	0.63	0.10	0.64	0.7	10.8	68.4	17.5	2.6	0.14	=.04	[0.01;0.28]
C	-0.10	0.78	-0.11	0.76	3.0	24.5	54.3	16.7	1.5	0.25	<.001	[0.13;0.36]
AC	-0.10	0.64	-0.16	0.80	1.5	33.1	48.0	14.5	3.0	0.14	=.02	[0.02;0.26]

Table S24

Participant's actual and perceived change of each preference dimension as well as the correlation in the sample of former participants of the Sociosexuality Study.

	actual change		perceived change		responses of perceived change (%)					insight		
	M	*SD*	*M*	*SD*	-2	-1	0	1	2	*r*	*p*	95% CI
SR	-.21	.72	.09	.72	1.54	13.85	61.54	20.00	3.08	.39	=.001	[.16; .58]
WT	.91	.94	.62	.76	0.00	3.08	46.15	36.92	13.85	.11	=.370	[-.13; .35]
VA	.16	.73	-.22	.67	1.54	29.23	60.00	7.69	1.54	.34	=.006	[.10; .54]
FO	.49	1.35	.72	1.07	6.15	1.54	32.31	33.85	26.15	.50	<.001	[.29; .66]
IQ	.94	.93	.34	.76	1.54	4.62	60.00	26.15	7.69	.12	=.347	[-.13; .35]
H	-	-	.28	.65	6.15	64.62	24.62	4.62	6.15	-	-	-
C	-.22	1.21	-.08	.82	4.62	21.54	52.31	20.00	1.54	.25	=.042	[.01; .47]
AC	.12	.98	-.14	.79	1.54	32.31	46.15	18.46	1.54	.25	=.048	[.00; .46]

S4. Conference Presentation

This research has been presented at the 21st Annual Meeting of the Society for Personality and Social Psychology in New Orleans (February 2020) and online at the International Association for Relationship Research (July, 2021).

S5. Author Contributions

JCD, LP and TMG planned T2 data collection. LP and TMG supervised the project. LP planned and collected T1 data. JDC collected T2 data, analysed T1 and T2 data and drafted the manuscript. All authors read the manuscript, provided critical revisions and approved the final manuscript.

S6. References

Borkenau, P., & Ostendorf, F. (1993). NEO-Fünf-Faktoren Inventar nach Costa und McCrae [NEO Five Factor Inventory after Costa and McCrae]. Göttingen, Germany: Hogrefe.

Lang, F. R., Lüdtke, O., & Asendorpf, J. B. (2001). Testgüte und psychometrische Äquivalenz der deutschen Version des Big Five Inventory (BFI) bei jungen, mittelalten und alten Erwachsenen. *Diagnostica*, *47*(3), 111-121. http://dx.doi.org/10.1026/0012-1924.47.3.111

Penke, L., & Asendorpf, J. B. (2008). Beyond global sociosexual orientations: A more differentiated look at sociosexuality and its effects on courtship and romantic relationships. *Journal of Personality and Social Psychology*, *95*(5), 1113–1135. https://doi.org/10.1037/0022-3514.95.5.1113

Appendix C.

Manuscript 3

Driebe, J.C. Stern, J., Penke, L. & Gerlach, T.M. (2021). *Ideals are Associated with Perceived Traits of Future Partners Across 13 Years, With a Closer Match Being Related to Commitment* [Manuscript in preparation]. Department of Biological Personality Psychology, University of Goettingen.

Ideals are Associated with Perceived Traits of Future Partners Across 13 Years, With a Closer Match Being Related to Commitment

Julie C. Driebe[1], Julia Stern[1,3], Lars Penke[1,2], Tanja M. Gerlach[1,2]

[1]University of Goettingen

[2]Leibniz Science Campus Primate Cognition

[3]University of Bremen

Author note

Correspondence concerning this article should be addressed to Julie Driebe. Email: juliedriebe@gmail.com

Abstract

Previous research is inconclusive on whether our ideal partner preferences guide relationship decisions. With the current study, we address this question using a longitudinal design across 13 years. We investigated whether ideal partner preferences are associated with perceived characteristics of partners and whether a closer match between ideals and the perception of a partner's traits is associated with better relationship outcomes (e.g., satisfaction or commitment). A sample of 178 participants (90 women, 88 men; $M = 45.7$ years, $SD = 7.2$) reported their ideal partner preferences in 2006. In 2019, they reported their relationship histories since the initial assessment, providing us with ratings of 322 relationships (39% of which are still ongoing). Using multilevel modelling, we found a positive association between participants' initially reported ideals and their perceptions of their partners' traits. Because the association was stronger with current ideals, preferences seem to be somewhat adjusted towards a partner. We operationalised the match between ideals and perceived partner traits using different metrics (e.g. Euclidean distance, corrected pattern metric). Across all matching operationalisations, we found that a closer match between ideals and perceived partner traits was associated with higher relationship commitment. Associations were more mixed for all other relationship outcomes. In line with the ideal standards model, preferences seem to be related to who we select as a partner and to how committed we are, whereas other models on human mate choice can also explain our pattern of results.

Keywords: ideal partner preferences, perceived partner traits, relationship outcomes

Humans spend an enormous amount of time and energy on finding the partner of their dreams (Fletcher et al., 2019), which is no wonder as romantic relationships exert a tremendous impact on our life's quality (Frost & Forrester, 2013; Holt-Lunstad et al., 2008; R. B. Miller et al., 2013; Pietromonaco & Collins, 2017). The high priority given to finding Mr. or Mrs. Right is also reflected in the literature: A bulk of research has investigated what we seek in an ideal partner (e.g. Buss, 1989; Buss & Barnes, 1986; Walter et al., 2020), operating under the (sometimes implicit) assumption that these preferences will guide mate choices (Campbell & Stanton, 2014). However, Campbell and Stanton (2014) emphasised the lack of studies actually investigating whether stated ideals indeed matched characteristics of a future partner and called for longitudinal designs with ideals being assessed prior to entering a relationship. To our knowledge, only two recent studies implemented such a design. Both studies tracked single participants over a period of five months and found ideals to be related to characteristics of future partners (Campbell et al., 2016; Gerlach, Arslan, et al., 2019). Yet, it still remains an open question whether ideals are related to attributes of future partners over a longer period of time. With the current study, we aimed to fill this gap: Employing unique data covering 13 years, we investigated whether ideal partner preferences are related to characteristics of future partners and whether this potential match between ideals and preferences is associated with a comprehensive set of relationship outcomes.

The Ideal Standards Model

Studies have found that, across cultures, humans share similar preferences for a desired partner (Buss, 1989; Walter et al., 2020). An ideal partner is often described as intelligent and kind (Buss & Barnes, 1986; Li et al., 2002), but preferences also typically differ between the sexes, with men placing higher importance on finding an attractive and healthy partner and women placing higher importance on finding a partner with a high status and resources (Buss & Schmitt, 2019; Feingold, 1990, 1992).

Despite these seemingly universal preferences, humans also differ from one another when it comes to their ideas about Mr or Mrs Right. The Ideal Standards Model (ISM) describes characteristics we desire in an ideal partner, often referred to as ideal standards or ideal partner preferences (Fletcher et al., 1999), and posits that they may differ between individuals (Campbell & Fletcher, 2015). While the ISM grouped these ideals into three main categories (warmth-trustworthiness, status-resources, vitality-attractiveness), other researchers have found additional dimensions such as confidence-humour, family orientation, or intellect (e.g. Csajbók & Berkics, 2017; Gerlach, Arslan, et al., 2019; Lam et al., 2016; Schwarz & Hassebrauck, 2012). The ISM proposes that these ideals are constantly accessible

and used to evaluate existing or potential partners and relationships by calculating a discrepancy between ideals and perceptions of partner traits (Campbell et al., 2001; Fletcher et al., 2000). Hence, a partner or relationship can be evaluated, the relationship and its dynamics (e.g., why conflicts arise) can be explained, and the relationship can be regulated, such that ideals or perceptions may be adjusted to reduce discrepancies between the two (for a detailed account, see Campbell et al., 2001; Fletcher et al., 2000).

The discrepancy between ideals and perceptions is presumed to be strongly related to our relationship's quality, above all relationship satisfaction (Fletcher et al., 1999; Fletcher et al., 2000; Overall et al., 2006). It is supposed to influence whether we are committed to a relationship (Rusbult et al., 2001) and whether we stay in the relationship or not (Fletcher et al., 2000). Further, romantic partners are both affected in their feelings of relationship satisfaction (Campbell et al., 2001; Campbell et al., 2013) and behaviours (e.g. focused on the prevention of negative outcomes, Lackenbauer & Campbell, 2012) as they are able to gauge whether they meet their partners' ideals. In sum, the ISM implies that our romantic relationships benefit from a closer match between our ideals and our partners' traits.

Yet, the ISM has been challenged in its proposition that ideals guide who we select as a romantic partner. In a meta-analysis, only in hypothetical scenarios were stated partner preferences associated with actual choices, whereas in face-to-face contexts no evidence for the predictive power of stated partner preferences was attained (Eastwick et al., 2014). The authors concluded that humans are simply not aware of what drives their mate choices and questioned the idea that preferences are related to actual choices (Eastwick et al., 2014).

Campbell and Stanton (2014) raised the concern that the studies available at that time only covered the initial stages of getting to know each other, but did not follow up on participants into longer relationships. In speed-dating studies, covering exactly these initial stages, physical attractiveness (Driebe et al., 2021; Hofer et al., 2021; Luo & Zhang, 2009; Todd et al., 2007) and characteristics easy to observe like age and height (Kurzban & Weeden, 2005) have been shown to be the main predictors of romantic interest. Similar results have been attained for mobile dating apps, with physical attractiveness again being the strongest predictor of users' romantic interest (Chopik & Johnson, 2021). Fletcher and colleagues (2014) and G. F. Miller and Todd (1998) suggested that because attractiveness is by far more easy to access compared to other characteristics, men and women alike could pay more attention to it in the very early stages of getting to know each other. These findings do not, however, negate the key ideas of the ISM, as ideals could still play a causal role at a later stage of relationship formation. Campbell and Stanton (2014) argued that, based on the

knowledge at the time, it was not possible to draw a final conclusion on the predictive validity of ideal partner preferences, or put differently: Whether humans select partners who more closely match their ideals and how this match is associated with their relationships' development. Studies that could actually speak to these issues would have to track participants over a longer period of time in which relationships are formed and assess ideals prior to entering a relationship. In the meanwhile, two studies implemented a research design and tracked single individuals over a period of five months. Campbell and colleagues (2016) found participants' initial preferences to be associated with their new partners' self-reported characteristics. Similarly, Gerlach, Arslan and colleagues (2019) found participants' initial preferences to be related with the perception of their future partners. While both studies are in line with the interpretation that we select partners who match our ideals, it still remains unclear what happens thereafter, limiting their evidence for hinting at a preference driven mate choice to a relatively short period of time. They also cannot speak to the implications preference-partner matching would have for key relationship outcomes, such as relationship quality. Previous studies that found a discrepancy between ideals and partner characteristics to be associated with relationship outcomes (e.g. the relationship's quality, or length of the relationship) investigated the link in already established relationships (Fletcher et al., 1999; Fletcher et al., 2000). Given that these studies assessed ideals and relationship outcomes at the same time, these studies cannot rule out that ideals were adjusted towards a partner. However, as Gerlach, Arslan and colleagues (2019) found ideals to be adjusted towards a partner when this partner fell short of participants' initial ideals, the possibility remains that happier relationships are the ones where individuals are more likely to change their ideals in order to match their partners. To sum up, there is still a need for longitudinal studies covering a longer time period with ideals assessed prior to individuals entering a romantic relationship and relationship outcomes thereafter.

The Present Work

With the current study, we aimed to close this gap in the literature by testing the ISM using a longitudinal study covering a time span of 13 years. We investigate the predictive validity of ideal partner preferences using several approaches, as each can yield unique insights (see Eastwick et al., 2019; Fletcher et al., 2020). First, as a first indicator of preferences' predictive validity, we investigate whether ideal partner preferences reported 13 years ago are associated with attributes of later partners. With attributes of later partners, we refer to participants' perception of their current and former partners since they have reported their ideal partner preferences. As outlined earlier, a positive association between ideals as

reported 13 years ago and participants' perception of their later partners (H1) would be consistent with a preference driven mate choice. While such a positive *ideal-trait-correlation* could still be the result of alternative explanations (e.g. a confounding third variable), a null-result would rule out a preference driven mate choice (Fletcher et al., 2020). Second, we investigate whether the match between ideals and a partner's traits, termed *ideal-partner matching* or *ideal partner preference-matching*[11] (Fletcher et al., 2020) is related to relationship outcomes. Because, so far, there is no consensus on the best strategy except that different approaches should be combined as they complement each other (see Eastwick et al., 2019; Fletcher et al., 2020), we operationalised this match using two different metrics: 1) the distance and the 2) correlation between ideals and perceived partner traits. We refer to both as an ideal partner preference-match and predict that a higher ideal partner preference-match is associated with better relationship outcomes (H2). The relationship outcomes investigated in the current study comprise a participant's self-reported relationship satisfaction, commitment, the length of a relationship, in case of relationship that have already ended, and who initiated a break-up. We predict a closer match between ideal partner preferences and a partner's traits to be associated with a higher relationship satisfaction, commitment and a longer relationship length. Regarding relationship dissolution, we predict that participants are less likely to initiate a break-up when there is a closer match between preferences and partner traits.

Finally, we aimed to gain new insights with an exploratory investigation: Because not everyone strives for being in a committed romantic relationship (Park et al., 2020), in an auxiliary analysis, we investigated how many individuals did not enter a romantic relationship across the 13 years study period and compared participants who entered a romantic relationship with those who did not across several characteristics (e.g. their personality traits, ideals or mate value).

Method

Sample

We recruited participants of a former study that took place in 2006 (Asendorpf et al., 2011), henceforth referred to as *T1*. Of the 382 initial participants, 226 participated in our online study (41% dropout) that was conducted in 2019, henceforth referred to as *T2*. We excluded 4 participants with a homosexual orientation because it was already an exclusion criterion at T1, where participants were explicitly asked about their partner preferences for

---

[11] These terms have recently evolved with earlier papers referring to it as *ideal-perception consistency* Fletcher et al. (2000) or simply *consistency* (Buyukcan-Tetik et al. (2017).

the opposite sex. Because we were interested in the association between preferences and partner characteristics, participants needed to have had at least one romantic relationship which exceeded six months in duration since the initial assessment or currently be involved in a romantic relationship. Therefore we excluded 25 participants who reported to have been single since T1 (11%). The remaining participants reported an overall of 362 relationships since T1, of which we had to exclude 26 relationships because participants either gave conflicting answers[12] ($n = 5$), indicated that the relationship had ended because of the death of the partner ($n = 2$) or specified the duration of the relationship as shorter than six months ($n = 19$) (see S2).

Our final sample consisted of 178 participants, comprising 90 women and 88 men with a mean age of $M = 45.7$ years ($SD = 7.2$, range = 31 - 66 years) and a total of 322 relationships. The majority (85%) of participants indicated a university degree as their highest level of education, the remaining 15% had some sort of school degree. Most participants were currently involved in a relationship (75%). Of the 322 relationships, 61% ($n = 196$) described former partners and 39% ($n = 126$) described a current partner. The mean age of all partners was $M = 44.3$ years ($SD = 8.5$, range = 21 - 81 years) with 165 (51%) being female and 157 (49%) being male. An attrition analysis revealed that participants who participated at T2 were less conscientious ($p = .049$, Hedges $g = -0.20$) and more neurotic ($p = .026$, Hedges $g = 0.23$) compared to participants who only participated at T1. No other significant group differences emerged (Table S2).

Because we expected to have a high dropout rate due to potentially outdated contact details, we also tried to re-contact former participants of a second study (Penke & Asendorpf, 2008b). Of these 142 initial participants, we were able to recruit another 66 for participation in a re-assessment. However, we were overly optimistic in being able to combine analyses of both samples. Initial assessment of ideal partner preferences diverged considerably between the two studies, making a combined analysis of both samples impossible in the end. Because of the more comprehensive assessment of ideal partner preferences and a way larger sample of the first study, we decided to provide analyses of this second sample in our supplement and interpret results of the first sample only (see S3 and for full results https://osf.io/tyc4r/?view_only=c0ea03dbbb004814b92426c3a31b9b3a).

Procedure

---

[12] Conflicting answers were for example if a participant initially reported to be single but later reported an ongoing relationship in the event-history calendar.

In 2006, single participants took part at a speed-dating event of a study called "Berlin Speed-Dating Study" (BSDS). Among other measures, participants' ideal partner preferences were assessed at the beginning of the event (for a detailed description, see Asendorpf et al., 2011).

From February 2019 to November 2019, we contacted these initial participants again and invited them to our online study. The study was implemented in the formr survey framework (formr.org; Arslan et al., 2020b), with the goal of investigating participants' romantic relationships longitudinally. As an incentive, participants received feedback on their personality and how their ideal partner preferences had changed over time. They also received a payment of 40€ when completing the study and a bonus of 10€ if they invited their peers and their partner to part in a separate part of the study. After being introduced to the study's goal, participants confirmed to have read the information about their data protection rights, the duration of the study and their incentive of taking part in the study and indicated their willingness to participate. Participants then filled out a short demographic questionnaire in which we assessed their age, gender and relationship status among other variables related to their personal life. Thereafter, participants reported their ideal partner preferences. As a next step participants filled out an event history calendar (EHC) in which we asked for participant's residences, jobs and important life events since their initial study participation (for the online implementation of the EHC, see Wieczorek et al., 2020). The calendar grid of the EHC served as a retrieval cue (Belli et al., 2001; Tully & Meyvis, 2017) to facilitate remembering all relationships since T1, with the goal to arrive at a complete reconstruction of participants' relationship histories. While presenting participants with their personalised EHC, they were asked to fill in all relationships exceeding six months since T1 as well as their current relationships independent of the relationships' length. Afterwards, we assessed more information on the relationships listed in the calendar. These variables included a partner's demographic information (e.g. age and gender), whether the relationship had ended and, if so, who initiated the break-up, how much the partner fulfilled certain characteristics and how satisfied and committed participants were in this relationship. Finally, participants filled out a number of personality measures, and provided further demographic information (e.g. on their professional life and living situation). Afterwards, participants had the possibility to invite up to two peers and, if applicable, their current partner to a second part of our study in which we strived to receive a self- and peer-rating of participant's partners.

Validation of Partner Ratings

In the second part of our study, we aimed to validate our focal participants' reports on their partners' traits with their partners' self-ratings. For participants who were currently involved in a romantic relationship and who agreed to invite their partner, we generated an anonymised link which they were asked to forward to their partners. Unfortunately, only 37 focal participants were interested in such a link and an even smaller number of partners participated. 12 partners ($n = 9$ women) with a mean age of 39.5 years (SD = 5.7, range = 32 - 51 years) rated themselves on 32 attributes. Focals' and partners' self-ratings were significantly associated ($b = 0.46$, 95% CI [0.36; 0.56], $p < .001$[13]). While we took this as indication for the validity of focals' ratings of partner traits, due to the small sample size, these results should of course be interpreted with caution (but see Gerlach, Arslan, et al., 2019, for a similar approach).

Since we were also interested in validating ratings of former partners, we asked our focal participants to invite up to two peers (e.g. friends or family members) who knew at least one of the partners mentioned in the EHC. For each peer, participants had the possibility to select which partner they would like their peer to rate. Thus, participants received up to two anonymised links which we asked them to forward to their peers. Again, unfortunately only 38 focal participants created at least one link. 19 peers ($n = 16$ women) with a mean age of 41.8 years (SD = 9.0, range = 28 - 58 years) participated in this second part of our study and rated 27 partners on 32 characteristics. Peer and focal ratings of partners were significantly associated ($b = 0.27$, 95% CI [0.19; 0.34], $p < .001$[14]).

Measures

Ideal Partner Preferences and Partner's Characteristics

At T1, participants rated a total of 59 items in their importance in an ideal partner on a 5-point Likert scale ranging from 1 (very unimportant) to 5 (very important). At T2, participants rated each partner as listed in the EHC on whether he or she possessed these characteristics on a scale from 1 (not at all) to 5 (very much). However, in order to reduce fatigue effects each partner was only rated on 25 instead of 59 characteristics. These 25 characteristics were previously selected in order to cover a broad variety of characteristics using a principal component analysis with oblimin rotation (see our preregistration on p.10 "2.1.4.Rating of (Former) Partners" for a detailed description

---

[13] In a multilevel model, we predicted focal participant's ratings of their partners with their partner's self-rating while including a random effect for each item and partner IDs.

[14] In a multilevel model, we regressed peer ratings on focal ratings of each partner while including a random intercept for each item, peer and partner IDs.

https://osf.io/x7rma/?view_only=cc43884bd1744047a8afb93106c42c0e) including the three dimensions used to describe an ideal partner as reported by Fletcher and colleagues (1999). Thus, we can only use the corresponding 25 items assessing participant's ideal partner preferences for our analyses (Table 1).

Table 1
Means and Standard Deviations of Perceived Partner Traits and Ideal Partner Preferences

	Traits		T1 Ideals	
	M	*SD*	*M*	*SD*
wealthy	2.62	1.07	2.11	0.92
financially secure	3.14	1.25	2.83	1.05
successful	3.18	1.10	2.79	0.92
understanding	3.64	1.05	4.13	0.68
sensitive	3.53	1.07	4.19	0.63
trustworthy	4.04	1.06	4.44	0.60
erotic	3.47	1.00	3.84	0.71
sexy	3.49	0.98	3.59	0.90
attractive	3.79	0.88	3.92	0.73
fond of children	3.85	1.16	3.68	1.08
family oriented	3.72	1.26	3.37	1.00
good mother/father	3.50	1.28	3.44	1.12
educated	3.75	0.98	4.21	0.67
acute	3.42	1.07	3.54	0.85
clever	3.38	0.99	3.48	0.93
funny	3.48	0.96	3.75	0.87
humorous	3.65	0.94	4.45	0.66
fun	3.52	0.94	3.85	0.81
shrewd	3.17	0.97	3.16	1.02
unconventional	3.10	1.11	3.20	0.90
imaginative	3.47	0.94	3.93	0.65
inventive	3.27	0.96	3.38	0.82
venturesome	2.77	1.16	2.72	0.89
assertive	3.41	1.07	3.59	0.76
confident	3.32	1.11	3.83	0.71

Note. Ideal partner preferences were assessed on a scale ranging from 1 (very unimportant) to 5 (very important). Perceived partner traits were assessed on a scale ranging from 1 (not at all) to 5 (very much).

Relationship Outcomes

For each partner, we assessed nine different relationship outcomes (Table 2). First, we assessed how satisfied participants were with the relationship on a scale ranging from 1 (not at all) to 5 (very). We tried to cover further aspects of relationship satisfaction with five additional single items. These aspects include sexual satisfaction, satisfaction with their standard of living, intention to plan a family, how much harmony they perceived in their relationship and how much they could rely on their partner. The tense varied in relation to a current and former partner. Each satisfaction score was analysed separately. For each current partner, we assessed participants' relationship commitment using three items inspired by Rusbult et al. (1998) on a scale from 1 (strongly disagree) to 5 (strongly agree). Analyses are based on the mean of the tree items (Cronbach's $\alpha = 0.86$, 95% CI [0.84; 0.89]). For each former partner, we calculated the relationship length based on participants' entry in the EHC. In cases where participants were in a relationship with the same person more than once, we summed up both durations. Additionally we assessed who initiated the break-up of the relationship on a scale from 1 (name of the former partner) to 5 (me).

Table 2
Item Content, Response Formats, Means and Standard Deviation of each Relationship Outcome

Outcome	Item [response format]	*M* (*SD*)
Satisfaction	How happy are/were you typically in your relationship with X? [1: not at all - 5: very]	3.39 (1.03)
	How harmonious is/was your relationship with X? [1: not at all - 5: very]	3.43 (1.11)
	Did/Do you imagine having a family with X? [1: not at all - 5: very]	3.18 (1.59)
	How passionate is/was your relationship with X? [1: not at all - 5: very]	3.28 (1.26)
	How satisfied are/were you with your standard of living with X? [1: not at all - 5: very]	3.74 (1.03)
	How much could you rely on X? [1: not at all - 5: very]	3.93 (1.18)
Commitment	"I'm oriented toward the long-term future of our relationship (e.g., I imagine being with my partner several	4.49 (0.82)

Outcome	Item [response format]	*M* (*SD*)
	years from now, I make plans for the future"	
	"I would feel very upset if our relationship were to end in the near future."	
	"I want our relationship to last for a very long time." [1: strongly disagree - 5: strongly agree]	
Relationship Length	Time between relationship start date (MM/YYYY) and end date (MM/YYYY) as entered in the EHC (in months). For still ongoing relationships, the we calculated the relationship length until the end of our data collection (11/2019)	57.79 (50.61)
Break-up	Who ended your relationship [1: X - 5: Me]	3.21 (1.61)

Note. For X we pasted partner names. We only assessed relationship commitment for current partners (*n* = 126) and break-up for former partners (*n* = 196).

(Semi-) Euclidean Distance

One of our two operationalisations to assess the ideal partner preference-match was to calculate the Euclidean distance as proposed by Conroy-Beam and Buss (2017). For each participant, we squared the difference between reported ideals and perceptions of a partner separately for each trait. We then calculated the square root of the sum of each squared difference. For each participant we received one value which we multiplied by minus 1 and added the maximum Euclidean distance plus one resulting in the following formula:

$$-1 * \sqrt{\sum_{1}^{n} (p_n - t_n)^2} + 21$$

With this transformation, we receive positive values in which higher scores represent a closer match between ideals and perceived partner traits. Because discrepancies might matter only if a partner falls short but not if a partner exceeds one's ideals (see Gerlach, Arslan, et al., 2019), we also calculated the Semi-Euclidean distance (Gerlach, Schultze, et al., 2019). For the Semi-Euclidean distance, we calculated the Euclidean distance as described above but only when the difference between traits and ideals fell below zero. Positive differences, which emerge when the perception of a partner's trait exceeds one's ideal, were set to zero.

(Partly) Corrected Pattern Metric

As a second operationalisation of the ideal partner preference-match, we calculated the corrected pattern metric (Eastwick et al., 2019). For each partner, we calculated the within-person correlation between participants' ideals and their perception of their partner using a Pearson product-moment correlation. We calculated this within-person correlation in two ways. First, we mean-centred participants' perceived partner ratings before correlating them with participants' ideals by calculating the across-sample mean of the partner ratings per item and subtracting this average from participants' corresponding perceived partner rating. This procedure removes the sample's average perception of a partner, the *normative profile* (Wood & Furr, 2016) and leaves the *distinctive profile* for each partner which expresses what characterises this person the most (Wood & Furr, 2016). This procedure has been proposed because the similarity between profiles is confounded with the desirability of certain characteristics. When removing the normative profile, this confounding variable is eliminated (for a detailed explanation, see Wood & Furr, 2016). Second, we also mean-centred participants' ideals following the same procedure as described above as they could also be confounded with the normative desirability of each characteristic (Wood & Furr, 2016). We then calculated the within-person correlation between participants' centred ideals and their centred partner perception. Mean-centring ideals and perception is referred to as the corrected pattern metric (Eastwick et al., 2019).

T2 Ideal Partner Preferences (Used for Robustness Checks)

At T2, participants rated the same 59[15] items like at T1 assessing their ideal partner preferences on the same scale. Again, only 25 of these items were used in our robustness checks because only these items corresponded to participants' ratings of how much they perceived their partners to fulfil these characteristics. One slight difference compared to the instructions at T1 was that at T2, we wanted to prevent ambiguities in our instructions and specified to rate a partner for a committed, long-term relationship. If participants were currently involved in a romantic relationship, it was noted to make each rating independently of one's current partner.

<div align="center">

Results

</div>

---

[15] At T2, participants rated 13 additional items assessing their ideal partner preferences which corresponded to the initially assessed ideal partner preferences of our second sample. Analyses involving these items can be found in the Supplement S3.

We analysed our data using R 4.0.2 (R Core Team, 2020a) and the packages lmerTest 3.1-2 (A. Kuznetsova et al., 2017), coxme 2.2-16 (Therneau, 2020), ordinal 2019.12-10 (Christensen, 2019), sjPlot 2.8.7 (D. Lüdecke, 2021) and ggplot2 3.3.3 (Wickham, 2016). We ran multiple robustness checks for each of our analyses. Detailed results of these robustness checks can be found on the OSF (https://osf.io/tyc4r/?view_only=c0ea03dbbb004814b92426c3a31b9b3a)

Only our study design but not all of our analyses can be regarded as pre-registered. We deviate from our pre-registered analyses because of two reasons: First, because it was not possible to combine both initial samples[16] and second, we do not think that all of our pre-registered analyses were reasonable[17]. We describe all deviations to our pre-registration including explanations for these differences in our supplement (S1).

Ideal-Trait Correlation

We fitted multilevel models in which we predicted participants' T1 ideal partner preferences with their perceived partner characteristics while including participants' age and gender and a random intercept for each characteristic as well as participant and partner IDs. In line with our hypothesis (H1), we found that ideal partner preferences were positively associated with perceived characteristics of future partners (Table 3). Controlling for effects of age and sex effects revealed that older participants gave their partner's lower ratings.

---

[16] For example, in our pre-registration we did not take into account that the Euclidean distance would necessarily differ between the two samples because of the diverging number assessing T1 ideal partner preferences, rendering a combined analysis impossible.

[17] For example, we analysed whether a match between ideals and perceived partner characteristics is associated with participants' relationship lengths. We initially intended to analyse this outcome only for ex-partners since the relationship lengths for ongoing relationships is unknown. However, later we decided that analysing this outcome using a Cox proportional hazards' mixed model is more reasonable because it allows us to include all relationships. The model calculates the relation of a variable (here the match between ideals and perceived characteristics) with the likelihood of an event (here the end of a relationship).

Table 3

Multilevel Model Predicting Perceived Partner Characteristics with Participant's T1 Ideals, Controlling for Age and Sex

Coefficient	Perceived Partner Characteristics				
	Estimates	std. Error	95% CI	Statistic	*p*-values
Intercept	3.43	0.07	3.30 – 3.56	52.04	<0.001
T1 Ideals (*z*-standardised)	0.17	0.01	0.15 – 0.20	12.88	<0.001
Age (*z*-standardised)	-0.11	0.03	-0.17 – -0.04	-3.30	0.001
Sex (0 = women, 1 = men)	0.03	0.07	-0.10 – 0.16	0.48	0.632

Note. We show only the relevant non-varying effects on the mean, see online materials for random effects.

Exploratorily, we then correlated each single ideal with the corresponding perceived partner trait. Descriptively, the majority of all correlations were positive, whereas the strength varied between different items. For example, there was a stronger association for attributes related to family orientation ($r = 0.21$), but a smaller association for attributes related to inventiveness ($r = 0.03$) (S2, Figure S2).

As robustness checks, we included an interaction between ideals and whether participants' were currently involved in a romantic relationship with this person. Further, we fitted the same model as an ordinal model because responses are more likely to be ordinal instead of interval scaled. Results of our robustness checks can be found on the OSF (https://osf.io/tyc4r/?view_only=c0ea03dbbb004814b92426c3a31b9b3a, "1_H1_ideal_trait_correlation.html"). Results replicated in both models: The association between ideals and partner characteristics was stronger when participants were currently involved in a romantic relationship with this person compared to an ex-partner (ideals: $b = 0.14$, 95% CI [0.11; 0.17], $p < 0.001$; current partner: $b = 0.44$, 95% CI [0.34; 0.54], $p < 0.001$; ideals * current partner: $b = 0.08$, 95% CI [0.04; 0.13], $p < 0.001$).

In addition, we predicted T2 ideal partner preferences with perceived partner characteristics and found that T2 ideals were also associated with perceived partner characteristics ($b = 0.26$, 95% CI [0.24; 0.28], $p < 0.001$). As indicated by the estimate and its confidence interval for T2 preferences and perceived partner traits, this association was stronger than association of T1 ideals with perceived partner traits. In sum, our results indicate a positive ideal-trait correlation.

Ideal Partner Preference-Matching

(Semi-) Euclidean Distance

We fitted multilevel models in which we predicted each of the nine relationship outcomes once with the Euclidean and once with the Semi-Euclidean distance while including participants' age and gender as additional predictors and a random intercept for the participant ID. One exception is participants' relationship length in which we used a Cox proportional hazards' mixed model (also known as survival analysis) instead of a multilevel model, though using the same predictors. This model allows us to investigate all relationships instead of excluding all still ongoing relationships because of their unclear length (see Footnote 7).

Results of the Euclidean and Semi-Euclidean distance models can be found in Table 4 (for full models and robustness checks see:

https://osf.io/tyc4r/?view_only=c0ea03dbbb004814b92426c3a31b9b3a,

1_H2_(semi)_euclidean_distance.html). The Euclidean distance was significantly associated with eight out of nine relationship outcomes and the Semi-Euclidean distance was significantly associated with all relationship outcomes. These results indicate that a lower distance between ideals and partner perceptions is associated with higher relationship satisfaction and commitment. In the case of the relationship length, we found that relationships lasted longer (i.e., reduced hazard rate) with a lower distance between ideals and perceived partner characteristics. When investigating who initiated a break-up, we found that partners instead of participants themselves were more likely to end the relationship with a lower distance between ideals and partner perceptions. However, this association was only significant for the Semi- and not the Euclidean distance.

Our pattern of results remained robust when including a variable on whether relationships were still ongoing compared to former relationships, although current relationships were rated more positively (see

https://osf.io/tyc4r/?view_only=c0ea03dbbb004814b92426c3a31b9b3a,

1_H2_(semi)_euclidean_distance.html). In addition, we determined how positive each partner was rated overall by calculating the mean perceptions of a partner's traits across all characteristics. When including this positivity score (Table 4), the Euclidean distance remained only significantly associated with three relationship outcomes (perceptions of harmony, intention to start a family and commitment) and the Semi-Euclidean distance remained only a significant predictor in four out of nine models (perceptions of harmony, intention to plan a family, reliability on a partner and commitment).

Results support our hypothesis (H2) suggesting that a higher ideal partner preference-match (operationalised as the Euclidean and Semi-Euclidean distance) is associated with

more positive relationship outcomes. However, this association seems to be partly due to how positive a partner is rated as the associations between relationship outcomes and our distance measures diminished when including the positivity score.

Table 4
Estimates and Confidence Intervals of the (Semi)-Euclidean Distance Predicting Each Relationship Outcome

Outcome	Estimate of the Euclidean distance [95% CI]		Estimate of the Semi-Euclidean distance [95% CI]	
	Main Model	with Positivity	Main Model	with Positivity
Satisfaction	0.31*** [0.20 – 0.42]	0.01 [-0.10 – 0.12]	0.55*** [0.45 – 0.65]	0.05 [-0.11 – 0.22]
Harmony	0.37*** [0.26 – 0.49]	0.19** [0.06 – 0.32]	0.48*** [0.37 – 0.60]	0.12** [0.06 – 0.32]
Family	0.54*** [0.38 – 0.70]	0.18* [0.01 – 0.35]	0.79*** [0.63 – 0.94]	0.30* [0.03 – 0.56]
Passion	0.23*** [0.10 – 0.37]	0.07 [-0.09 – 0.22]	0.12*** [0.06 – 0.18]	-0.09 [-0.32 – 0.14]
Living	0.25*** [0.14 – 0.36]	0.03 [-0.09 – 0.15]	0.41*** [0.30 – 0.51]	0.08 [-0.11 – 0.26]
Reliability	0.38*** [0.26 – 0.50]	0.13 [-0.00 – 0.26]	0.56*** [0.44 – 0.68]	0.23* [0.03 – 0.42]
Commitment	0.28*** [0.11 – 0.45]	0.21* [0.04 – 0.38]	0.44*** [0.27 – 0.61]	0.42** [0.16 – 0.69]
Relationship Length	-0.30*** [-0.43 – -0.16]	0.02 [-0.16 – 0.20]	-0.46*** [-0.60 – -0.33]	-0.14 [-0.40 – 0.13]
Break-up	0.21 [-0.43 – 0.01]	-0.02 [-0.26 – 0.22]	-0.35** [-0.59 – -0.12]	-0.03 [-0.41 – 0.35]

Note. *** $p < .001$, ** $p < .01$, * $p < .05$. [95% CI] = 95% confidence intervals. Distance scores were *z*-standardised before predicting each relationship outcome. In the column "with positivity", the estimate of the (semi)-Euclidean distance on each relationship outcome is listed while controlling the model on how positive each partner was rated. Full models can be found at: https://osf.io/tyc4r/?view_only=c0ea03dbbb004814b92426c3a31b9b3a

(Partly) Corrected Pattern Metric

We then fitted multilevel models for each relationship outcome in which we predicted each outcome with the within-person correlation[18] as well as a random intercept for the participant ID. We did so separately for both calculated within-person correlations. When predicting the relationship length we used a Cox proportional hazards' mixed model (survival analysis) instead of a multilevel model, though again using the same predictors.

Results can be found in Table 5 (for full models and robustness checks see: https://osf.io/tyc4r/?view_only=c0ea03dbbb004814b92426c3a31b9b3a, 1_H2_(partly)_corrected_pattern_metric.html). When predicting relationship outcomes with the partly centred within-person correlation (only partner perceptions but not ideals centred), we found that correlation coefficients were significantly associated with six out of nine relationship outcomes. This means for example that with a higher correlation between ideals and perceived partner traits, participants rated their relationship as more satisfying. More specifically, satisfaction ratings increased from a value of about 2.8 when the correlation was -0.60 to about 3.70 when it was 0.60 (Figure 2).

---

[18] Previous research Fisher z-transformed each correlation coefficient for a more normal distribution. We did not do so because correlation coefficients were already normally distributed and predictors in a regression model do not need to follow a normal distribution (Field et al., 2014).

Figure 2

*Association between Relationship Satisfaction and the Correlation between Ideals
(uncentred) and Perceived Partner Traits (mean-centred)*
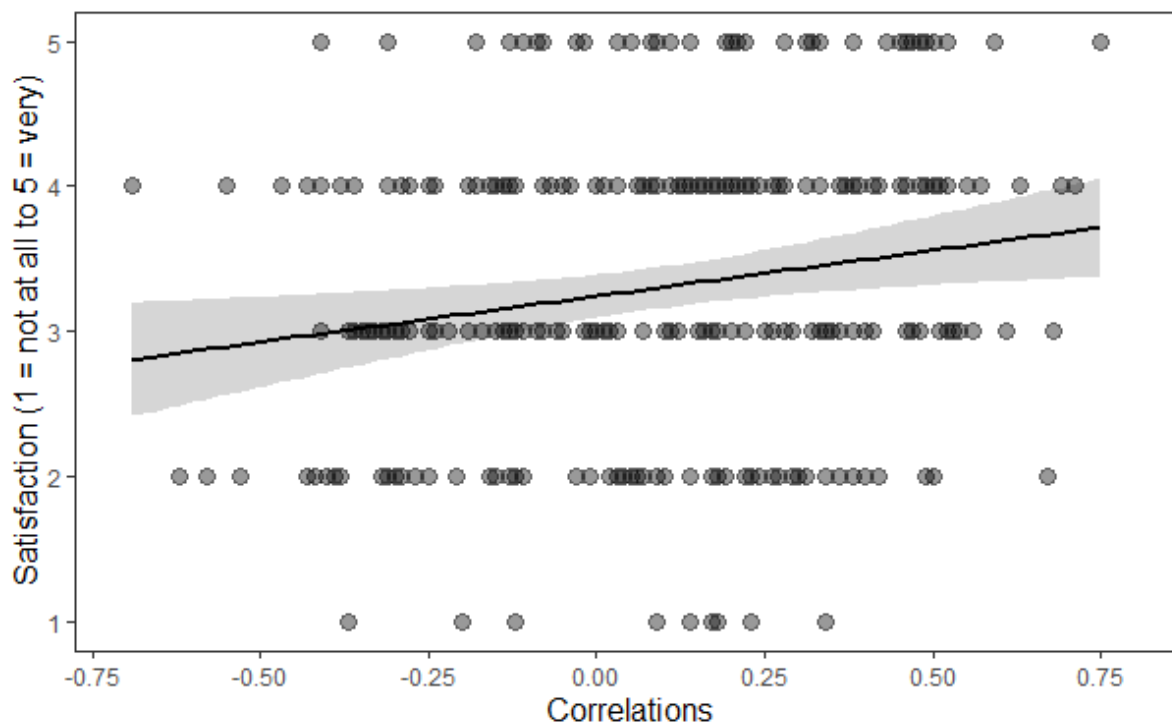
Note. The figure demonstrates that a higher correlation between ideals and perceived partner
characteristics is associated with a higher relationship satisfaction. The solid line illustrates
the regression line with the shaded area in grey representing the 95% confidence interval. The
dots represent the participant's correlation coefficients whereas varying opacity is caused by
multiple participants overlapping.

However, when predicting relationship outcomes with the corrected pattern metric
(correlation of centred partner perceptions and centred ideals), correlation coefficients were
significantly associated with only two out of nine relationship outcomes (satisfaction with the
standard of living and commitment).

Table 5

Estimates of the (Partly) Corrected Pattern Metric Predicting Each Relationship Outcome

Outcome	partly corrected pattern metric	corrected pattern metric
Satisfaction	0.65*** [0.27 – 1.03]	0.10 [-0.33 – 0.52]
Harmony	0.96*** [0.56 – 1.37]	0.22 [-0.24 – 0.67]
Family	0.91**	0.28

Outcome	partly corrected pattern metric	corrected pattern metric
	[0.31 – 1.50]	[-0.37 – 0.94]
Passion	0.21 [-0.26 – 0.69]	-0.06 [-0.58 – 0.47]
Living	0.14 [-0.26 – 0.53]	0.44* [0.01 – 0.87]
Reliability	1.04*** [0.61 – 1.48]	-0.37 [-0.12 – 0.85]
Commitment	1.05** [0.52 – 1.57]	0.58* [0.03 – 1.13]
Relationship Length	-0.84** [-1.35 – -0.32]	-0.21 [-0.78 – 0.36]
Break-up	0.18 [-0.57 – 0.93]	-0.23 [-1.05 – 0.59]

Note. *** $p < .001$, ** $p < .01$, * $p < .05$. [95% CI] = 95% confidence intervals. Full models can be found at: https://osf.io/tyc4r/?view_only=c0ea03dbbb004814b92426c3a31b9b3a

In our robustness checks, we included participant's age and sex as control variables. Further, we repeated our analyses calculating the within-person correlation as Kendall rank coefficients and fitted cumulative logit link (i.e. ordinal) mixed models because responses are more likely to be ordinal instead of interval scaled. Our previously described results remained virtually identical across all robustness checks.

Overall, these results suggest that when we correct for the normative desirability of traits, there seems to be an association for some relationship outcomes and the match between ideals and a partner's traits. However when we also norm-correct ideals, this association reduces dramatically which makes results supporting our second hypotheses more mixed.

Table 6

Summary of Significant Effects for Relationship Outcomes across Analyses

Sig. effects for the outcome ...	Analyses					
	Euclidean Distance		Semi-Euclidean Distance		Partly corrected pattern metric	Corrected pattern metric
	Main Model	with Positivity	Main Model	with Positivity		
Satisfaction	Yes	-	Yes	-	Yes	-
Harmony	Yes	Yes	Yes	Yes	Yes	-
Family	Yes	Yes	Yes	Yes	Yes	-
Passion	Yes	-	Yes	-	-	-
Living	Yes	-	Yes	-	-	Yes
Reliability	Yes	-	Yes	Yes	Yes	-
Commitment	Yes	Yes	Yes	Yes	Yes	Yes
Relationship Length	Yes	-	Yes	-	Yes	-
Break-up	-	-	Yes	-	-	-

Note. The "Yes" in a cell indicates that the estimate of the match between ideals and perceived partner characteristics was significant, whereas the "-" indicates a no significant estimate.

Long-Term Singles

When calculating the corrected pattern metric, the normative desirability of an item is calculated by averaging each item across the whole sample. However, this aggregation suffers from the problem that the normative profile is only representative of participants who were at least in one romantic relationship since T1, thus, broadly speaking of participants who were selected as a partner. However, participants without any relationship since T1 are not included in this normative profile. Therefore, in exploratory analyses, we tested whether participants who have not entered a romantic relationship since T1 (referred to as long-term singles) differed to participants with at least one relationship (referred to as final sample) in self-reported personality traits and demographic variables.

25 participants (*n* = 12 women) reported not to have had a relationship lasting longer than 6 months since T1 and were also currently not in a committed relationship. We compared these two groups using independent *t* tests. Results can be found in Table 7. Because of the small size of the long-term singles group this comparison should be interpreted with caution. We found that long-term singles were on average older, were searching for a short-term partner to a lower extent, had a lower self-perceived mate value, and were more open to experiences compared to our final sample. We found no other significant group differences.

Descriptively, both groups had similar high ideal partner preferences. But for long-term singles, ideals decreased over time whereas our final sample had similarly high (descriptively even slightly increased) ideal partner preferences at T2. Further, descriptively, long-term singles rated themselves to fulfil ideal partner characteristics to a smaller extent as compared to our final sample.

Table 7
Comparison of Participants With to Participants Without Committed Relationships Since T1

Variable	response scale	Final Sample (*n* = 178)	Long-term singles (*n*= 25)	*p*-value	Hedges *g*
		M (*SD*)	*M* (*SD*)		
Age	in years	45.72 (7.24)	49.04 (7.46)	=0.045	0.46
Sexual Orientation	1 (*homosexual*) - 7 (*heterosexual*)	1.08 (0.32)	1.08 (0.28)	=0.944	-0.01
Years of Education	in years	16.49 (7.22)	17.16 (6.94)	=0.658	0.09
search for a long-term partner	1 (*not searching at all*) - 7 (*searching a lot*)	3.62 (1.84)	3.24 (1.71)	=0.388	-0.21
search for a short-term partner	1 (*not searching at all*) - 7 (*searching a lot*)	2.53 (1.67)	1.76 (1.23)	=0.031	-0.50
Openness	1 (*does not apply to me at all*) - 5 (*applies to me a lot*)	3.90 (0.58)	4.20 (0.61)	=0.026	0.52
Conscientiousness	1 (*does not apply to me at all*) - 5 (*applies to me a lot*)	3.64 (0.62)	3.57 (0.68)	=0.626	-0.11
Extraversion	1 (*does not apply to me at all*) - 5 (*applies to me a lot*)	3.44 (0.75)	3.29 (0.84)	=0.382	-0.21
Agreeableness	1 (*does not apply to me at all*) - 5 (*applies to me a lot*)	3.54 (0.60)	3.44 (0.64)	=0.448	-0.17
Neuroticism	1 (*does not apply to me at all*) - 5 (*applies to me a lot*)	2.85 (0.84)	2.90 (0.90)	=0.799	0.06
Sociosexual desire	1 (*never*) - 5 (*nearly every day*)	2.70 (1.11)	2.67 (0.99)	=0.863	-0.03

Variable	response scale	Final Sample (*n* = 178)	Long-term singles (*n*= 25)	*p*-value	Hedges *g*
		M (*SD*)	*M* (*SD*)		
Sociosexual attitude	1 (*not at all*) - 5 (*very much*)	3.23 (1.17)	3.21 (1.14)	=0.939	-0.02
Sociosexual behaviour	no. coded	3.50 (0.76)	3.60 (1.23)	=0.862	0.13
No. of life-time sexual partners	no.	23.90 (38.58)	23.32 (26.53)	=0.928	-0.20
Mate Value	1 (*not at all*) - 5 (*very much*)	3.26 (0.86)	2.60 (1.22)	=0.014	-0.73
T1 ideals	1 (*very unimportant*) - 5 (*very important*)	3.48 (0.31)	3.48 (0.35)	=0.986	0.00
T2 ideals	1 (*very unimportant*) - 5 (*very important*)	3.51 (0.40)	3.40 (0.43)	=0.228	-0.28
self-rating on ideals	1 (*not at all*) - 5 (*very much*)	3.73 (0.39)	3.55 (0.46)	=0.065	-0.46
facial attractiveness (ratings by 15 opposite-sex same-aged at T1)	1 (*not attractive at all*) - 7 (*very attractive*)	2.87 (0.91)	2.70 (0.78)	=0.335	-0.19

Note. Only participants who reported to be currently single reported their search for a short-term and long-term partner. Because the items assessing sociosexual behaviour and their total number of sexual partners were non-compulsory, not all participants responded to these items.

Discussion

In the current study, drawing on unique data across 13 years, we investigated whether participants' ideal partner preferences were associated with perceived partner characteristics of their future partners (ideal-trait correlations) and whether a closer match between ideals and perceived partner characteristics was associated with various relationship outcomes such as a person's satisfaction or commitment (ideal partner preference-matching).

Supported Ideal-Trait Correlation

Across a time span of 13 years, we found that initially reported preferences were indeed associated with participants' perception of their future partners, supporting our first hypothesis (H1). Current ideal partner preferences showed a stronger association with the perception of a partner's traits compared to the initially reported preferences, a pattern consistent with the idea that ideals may be somewhat malleable. In particular, it might be that ideals get adjusted to match a partner's traits (Conroy-Beam & Buss, 2016; Fletcher et al., 2000), especially when a partner falls short of initial expectations (Gerlach, Arslan, et al., 2019). As such, our results are in line with both, an account of preference driven mate choice

and a post-hoc update of preferences (see Conroy-Beam & Buss, 2016). However, we cannot rule out some alternative explanations for this positive association, such as, for example a confounding third-variable (Fletcher et al., 2020) or individuals spending their lives in environments where they encounter a number of well-matching "potentials", but choice within these environments being more or less random (Gerlach, Arslan, et al., 2019). Hence, the positive ideal-trait correlation attained in the present study should not be interpreted in a causal way, even though a null-finding would clearly have ruled out a preference driven mate choice (Fletcher et al., 2020).

Partly Supported Ideal Partner Preference-Matching

We then investigated whether a closer match between ideals and preferences was associated with better relationship outcomes. Our findings partly supported our second hypothesis (H2). Across all analyses, the outcome relationship commitment was associated with a match between ideals and perceived partner characteristics. However, for all other relationship outcomes results were more mixed and varied with our analytic choices (see Table 6). Across most analyses, the outcome of how harmonious the relationship was perceived, how much participants relied on their partner, and whether participants imagined having a family with their partner were associated with a closer match between ideals and perceived partner characteristics. The results for all other outcomes were more mixed, with the weakest evidence for the outcome break-up (see Table 6).

Regarding the analytic choices, results of the Euclidean distance seemed straightforward: a smaller distance between ideals and perceptions of a partner's traits was associated with stronger commitment, longer relationships and higher relationship satisfaction, even covering various domains of satisfaction. When we only determined the distance for characteristics where a partner fell short but not exceeded ideals (the semi-Euclidean distance), the associations were even stronger and we also found that with a smaller distance, participants were less likely to be the person who ended the relationship. Participants being more sensitive to partners falling short compared to exceeding ideals, supports previous research (Buyukcan-Tetik et al., 2017). However, these associations reduced or even diminished when we included how positive a partner was rated in general, allowing several interpretations. One explanation could be that the match between ideals and traits is simply not related to whether a person is, for example, happy in a relationship, but that seeing the partner in a positive light more generally is more decisive when it comes to happiness in a relationship.

But why might commitment as a relationship outcome be less affected by such a general positivity effect? In comparison to the other outcomes, the assessment of commitment might have been more reliable as it was the only outcome which was not assessed with a single item but with three items. Another difference of this outcome was that it was only assessed for current relationships since the items' content covered the feelings and orientation towards a continuing of the relationship which would not have made sense to be assessed for ex-partners. In all other models but this, the positivity score could have been confounded with whether participants rated a current or a former partner, since still ongoing relationships compared to ex-partners were rated more positively. Therefore, the positivity score is difficult to disentangle with the influence of whether a relationship is still ongoing. Further, including a positivity score might be too conservative because it could entail the reason why people are in a relationship with that person, namely because they see their partner favourably. Finally, a positivity score might not even be necessary as our results and also previous literature suggests that participants have an accurate perception of their partners (e.g. Allik et al., 2016; Connelly & Ones, 2010; Watson et al., 2000), even while being positively biased (Fletcher et al., 2014).

We also operationalised the match between ideals and preferences using the (partly) corrected pattern metric, a metric that has the benefit of not being confounded with mean level effects of positivity in judgments (Fletcher et al., 2020). Our pattern of results, however, remained ambiguous: The partly corrected pattern metric suggested a positive association for most relationship outcomes (6/9), whereas the corrected pattern metric found a significant association for a minority of outcomes (2/9).

Both operationalisations remove the normative perceptions of a partner's traits (i.e., the average profile of a partner in the sample), whereas the corrected pattern metric additionally removes the normative component in participants' ideals (i.e., the average preference profile regarding an ideal partner in our sample). Norm-correcting perceptions of a partner's traits and ideals may remove the social desirability of traits as a confound, which especially affects highly desirable items (Wood & Furr, 2016). What remains is the distinctive profile of a partner, respectively participants' ideals, which distinguishes partners/ideals from other people (Wood & Furr, 2016). Wood and Furr (2016) argue that almost all variables are affected by a normative-desirability confound because they found that next to personality perceptions of others also variables such as emotions, attitudes and values were affected. Based on these findings, Eastwick and colleagues (2019) argued that participants' ideal partner preferences and their perceptions of their partners' characteristics

need to be corrected for this confound. But because no study investigated so far, whether ideal partner preferences really are affected, we repeated our analyses once with and once without norm-correcting ideals.

This correction is conservative (Wood & Furr, 2016). Funder (2001) argues that it removes valid variance since it is reasonable to partly receive an accurate perception of another person because of a knowledge of what an average person is like. Because we only assessed the relationship outcomes with single items, removing valid variance could particularly affect our results. Again commitment, the only outcome we assessed with three items, still had a positive, although reduced, association with our measures operationalising the match between ideals and a partner's traits. Another limiting factor of this measure is that normativeness is determined as the average perception of a partner in our sample. Thus, it reflects only profiles of participants who have been in at least one relationship in the investigated time span. Yet, 14% of our sample have not entered a relationship during the 13 years that we investigated. Although this sample is too small in order to draw strong conclusions, our analyses suggest the possibility that the personality or demographic profiles of these participants could be different as compared to participants who had been in at least one relationship during the respective time span. The normative profile could therefore be not representative of a profile of an average person but only of an average partner, thus, not reflecting people who are never chosen as a romantic partner. In contrast, the aim of controlling for a normative profile is to control for participants' perceptions of what an average person, not what an average partner, is like. Thus, correcting partner perceptions using the average rating across only participants who were in a relationship, thus selected as a partner, might be too conservative or even not representative. Future studies are therefore needed which investigate this potential confounding more closely.

Our Results Replicate Previous Research

Overall, our results suggest that participants have a higher commitment to their relationship when there is a closer match between their ideals and their partners' traits, which is in line with previous studies (Rusbult et al., 2001). Our results are less clear for the other relationship outcomes. One possibility for these mixed findings is that our single item measures of the other relationship outcomes were not reliable enough to uncover a link. Studies finding associations mostly used scales to measure relationship outcomes, for example Fletcher and colleagues (1999) and Fletcher and colleagues (2000) used a 6-item measure assessing the relationship's quality. Another possibility is that commitment has the strongest association with the match between ideals and preferences because it could plays a

more central role in the functioning of a relationship compared to the other relationship outcomes: Rusbult's investment model (2001) pronounces the vast importance of commitment for a relationship's quality as it promotes cognitions and behaviours maintaining the relationship and thus, influences participants' satisfaction. In line with this claim, Joel and colleagues (2020) have found that commitment was a central predictor of relationship quality.

Despite the possibility that commitment is more central compared to other relationship outcomes, diverging measurements could also explain the mixed findings. For example, studies directly asking participants in how far their partners matched their ideals found associations with their marital satisfaction (Campbell et al., 2013), relationship quality (Campbell et al., 2013; Overall et al., 2006) and romantic interest in a speed-dating design (Fletcher et al., 2014). Studies using the pattern metric (neither norm-correcting ideals nor perceptions) found associations between the ideal partner preference-match and relationship outcomes (e.g. Buyukcan-Tetik et al., 2017; Eastwick & Neff, 2012; Fletcher et al., 1999; Fletcher et al., 2000). Our mixed results when using the corrected pattern metric replicate previous research. To our knowledge, only three studies used the corrected pattern metric with one study finding an association (Fletcher et al., 2020), one finding no association (Eastwick et al., 2019), and one study finding an association in a Taiwan but not US sample (Lam et al., 2016). Although in our opinion we should be able to find an effect independent of measurement, the corrected pattern metric might be too conservative as it suffers under the limitation that "normativeness" is limited to participants who entered relationships.

Do preferences guide who we select as a partner?

Despite these mixed findings, the overall pattern of results is in line with the ISM as what people want in a partner seems to be existent in their future partners. Further, the discrepancy between ideals and a partner's traits influences, at least to some parts, the evaluation and regulation of the relationship. But apart from the ISM, other models on human mate choice are also able to explain our pattern of results. For example, we only know with whom participants entered a relationship, but not who they might have rejected as a partner. Perhaps another person needed to exceed a certain threshold of a characteristic in order to be considered as a partner (e.g. this person needs to exceed a minimum level of attractiveness in order to be considered for a next step of relationship formation) as proposed by the sequential-aspiration threshold model (Todd & Miller, 1999). There could even be several models at work (G. F. Miller & Todd, 1998): For example, at the beginning of relationship formation another person needs to exceed a certain threshold of physical attractiveness in order to be still viewed as a potential partner. When participants get to know this other person

more closely, they form a more accurate perception of the other person, for example, of the reliability of that person. At this later stage of relationship formation, this person could then be compared to how important participants rate the reliability of a partner. Only if the other person's reliability is close to this ideal expectation, the participant might choose to enter a relationship with this specific person. This could explain why at initial encounters and in speed-dating studies, characteristics which are easy to observe (e.g. physical attractiveness) drive mate choices, but at later steps of relationship formation, potential partners might rather be compared to ideals (Fletcher et al., 2014).

Strengths

Our results show that ideal partner preferences are associated with the selection of a partner across a long period of time since we found that ideals were associated with perceived partner characteristics of partners found in the subsequent 13 years and since we found a closer match between the two was related to a higher relationship commitment. Because our findings hold using several statistical approaches, our findings can contribute to the current scientific discourse in which the influence of preferences on the selection of a partner was questioned and previous analyses have been debated (e.g. Eastwick et al., 2019; Fletcher et al., 2020). Further, our longitudinal design assessing ideals when participants were single and their relationship history thereafter covering a period of 13 years, makes our study data very unique compared to previous studies: First, because to our knowledge, no study investigated such a long time span and second, because many studies assessed participants' ideals simultaneously with how they perceived their partners to fulfil these ideals. These designs cannot preclude that reported ideals were adjusted in order to match their partners and thus, whether ideals guide the selection of a partner. Next to these strengths, our older and more diverse sample with more relationship experiences, could explain why we, compared to other studies (e.g. Eastwick et al., 2019; but see Fletcher et al., 2020), found evidence for preferences influencing human mate choice.

Limitations and Future Research

Our study did not come without limitations. First, although we recruited a more diverse sample with regard to a higher age range and a more equal distribution of both sexes in comparison to previous studies, our sample is still highly educated and from Western background, limiting the representativity of our results. Future studies should aim to recruit more representative samples, including participants from non-Western countries (Henrich et al., 2010b). Second, we assessed all relationship outcomes except for relationship commitment with a single item. This less reliable assessment could explain the inconclusive

results for all other relationship outcomes. Third, although our data allows us to investigate a time span of 13 years, more than two assessment time points would have been preferable. Now, ratings of former partners were made retrospective, making potential biases more likely. Despite our efforts to validate participants' ratings of their partners with reports of others, only few peers and partners themselves participated in our study. However, the ratings of these peers and partners were correlated with our focal ratings suggesting that perceived partner ratings are valid. Future studies similar to Eastwick and colleagues (2021) which track participants' networks across several assessments to investigate the different stages of relationship development while also taking into account who might be rejected as a partner, can shed further light on human mate choice. Nevertheless, future research should also investigate more closely what differs between people who enter versus those who do not enter romantic relationships over a long time span. Finally, as aforementioned, we cannot explain the exact mechanisms of how ideal partner preferences are integrated into mate choices. As suggested by Conroy-Beam (2021) and Brandner et al. (2020), we encourage future research to test specific models of how preferences are integrated into human mate choices (e.g. the sequential-aspiration threshold model).

Conclusion

In conclusion, we found that ideal partner preferences are associated with the perception of a future partner's traits, despite ideals being somewhat adjusted towards a partner. Across all measures, we found that participants reported a higher commitment towards their relationship when there was a higher match between their ideals and their perceptions of their partner. For all other relationship outcomes, this association varied with our analytic choices, highlighting the need for future research which tracks single individuals over long periods of time including several assessments of their preferences, tracking of participants' social networks from which they may select their partners and a more comprehensive measurement of different relationship outcomes.

References

Allik, J., Vries, R. E. de, & Realo, A. (2016). Why are moderators of self-other agreement difficult to establish? *Journal of Research in Personality*, *63*, 72–83. https://doi.org/10.1016/j.jrp.2016.05.013

Arslan, R. C., Walther, M. P., & Tata, C. S. (2020). formr: A study framework allowing for automated feedback generation and complex longitudinal experience-sampling studies using R. *Behavior Research Methods*, *52*(1), 376–387. https://doi.org/10.3758/s13428-019-01236-y

Asendorpf, J. B., Penke, L., & Back, M. D. (2011). From Dating to Mating and Relating: Predictors of Initial and Long–Term Outcomes of Speed–Dating in a Community Sample. *European Journal of Personality*, *25*(1), 16–30. https://doi.org/10.1002/per.768

Belli, R. F., Shay, W. L., & Stafford, F. P. (2001). Event History Calendars and Question List Surveys: A Direct Comparison of Interviewing Methods *Public Opinion Quarterly*, *65*(1), 45–74. https://doi.org/10.1086/320037

Brandner, J. L., Brase, G. L., & Huxman, S. A. (2020). "Weighting" to find the right person: compensatory trait integrating versus alternative models to assess mate value. *Evolution and Human Behavior*, *41*(4), 284–292. https://doi.org/10.1016/j.evolhumbehav.2020.05.001

Buss, D. M. (1989). Sex differences in human mate preferences: Evolutionary hypotheses tested in 37 cultures. *Behavioral and Brain Sciences*, *12*(1), 1–14. https://doi.org/10.1017/S0140525X00023992

Buss, D. M., & Barnes, M. (1986). Preferences in human mate selection. *Journal of Personality and Social Psychology*, *50*(3), 559–570. https://doi.org/10.1037/0022-3514.50.3.559

Buss, D. M., & Schmitt, D. P. (2019). Mate Preferences and Their Behavioral Manifestations. *Annual Review of Psychology*, *70*(1), 77–110. https://doi.org/10.1146/annurev-psych-010418-103408

Buyukcan-Tetik, A., Campbell, L., Finkenauer, C., Karremans, J. C., & Kappen, G. (2017). Ideal Standards, Acceptance, and Relationship Satisfaction: Latitudes of Differential Effects. *Frontiers in Psychology*, *8*, 1691. https://doi.org/10.3389/fpsyg.2017.01691

Campbell, L., Chin, K., & Stanton, S. C. E. (2016). Initial Evidence that Individuals Form New Relationships with Partners that More Closely Match their Ideal Preferences. *Collabra*, *2*(1). https://doi.org/10.1525/collabra.24

Campbell, L., & Fletcher, G. J. O. (2015). Romantic relationships, ideal standards, and mate selection. *Current Opinion in Psychology*, *1*, 97–100. https://doi.org/10.1016/j.copsyc.2015.01.007

Campbell, L., Overall, N. C., Rubin, H., & Lackenbauer, S. D. (2013). Inferring a partner's ideal discrepancies: Accuracy, projection, and the communicative role of interpersonal behavior, *105*(2), 217–233. https://doi.org/10.1037/a0033009

Campbell, L., Simpson, J. A., Kashy, D. A., & Fletcher, G. J. O. (2001). Ideal Standards, the Self, and Flexibility of Ideals in Close Relationships. *Personality and Social Psychology Bulletin*, *27*(4), 447–462. https://doi.org/10.1177/0146167201274006

Campbell, L., & Stanton, S. C. E. (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*(9), 485–494. https://doi.org/10.1111/spc3.12126

Chopik, W. J., & Johnson, D. J. (2021). Modeling dating decisions in a mock swiping paradigm: An examination of participant and target characteristics. *Journal of Research in Personality*, *92*, 104076. https://doi.org/10.1016/j.jrp.2021.104076

Christensen, R [R.H.B.]. (2019). *ordinal- Regression Models for Ordinal Data* (Version 2019.12-10) [R package]. https://CRAN.R-project.org/package=ordinal

Connelly, B. S., & Ones, D. S. (2010). An other perspective on personality: Meta-analytic integration of observers' accuracy and predictive validity. *Psychological Bulletin*, *136*(6), 1092–1122. https://doi.org/10.1037/a0021212

Conroy-Beam, D. (2021). Couple Simulation: A Novel Approach for Evaluating Models of Human Mate Choice. *Personality and Social Psychology Review*, *25*(3), 191–228. https://doi.org/10.1177/1088868320971258

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*(1), 53–66. https://doi.org/10.1037/pspi0000054

Conroy-Beam, D., & Buss, D. M. (2017). Euclidean distances discriminatively predict short-term and long-term attraction to potential mates. *Evolution and Human Behavior*, *38*(4), 442–450. https://doi.org/10.1016/j.evolhumbehav.2017.04.004

Csajbók, Z., & Berkics, M. (2017). Factor, factor, on the whole, who's the best fitting of all? Factors of mate preferences in a large sample. *Personality and Individual Differences*, *114*, 92–102. https://doi.org/10.1016/j.paid.2017.03.044

Driebe, J. C., Sidari, M. J., Dufner, M., Heiden, J. M. von der, Bürkner, P. C., Penke, L., Zietsch, B. P., & Arslan, R. C. (2021). Intelligence can be detected but is not found attractive in videos and live interactions. *Evolution and Human Behavior.* Advance online publication. https://doi.org/10.1016/j.evolhumbehav.2021.05.002

Eastwick, P. W., Finkel, E. J., & Simpson, J. A. (2019). Best Practices for Testing the Predictive Validity of Ideal Partner Preference-Matching. *Personality and Social Psychology Bulletin*, *45*(2), 167–181. https://doi.org/10.1177/0146167218780689

Eastwick, P. W., Joel, S., Molden, D. C., Finkel, E. J., & Carswell, K. L. (2021). *Predicting Romantic Interest during Early Relationship Development: A Preregistered Investigation using Machine Learning* [Manuscript submitted for publication]. University of California, Davis.

Eastwick, P. W., Luchies, L. B., Finkel, E. J., & Hunt, L. L. (2014). The predictive validity of ideal partner preferences: A review and meta-analysis. *Psychological Bulletin*, *140*(3), 623–665. https://doi.org/10.1037/a0032432

Eastwick, P. W., & Neff, L. A. (2012). Do Ideal Partner Preferences Predict Divorce? A Tale of Two Metrics. *Social Psychological and Personality Science*, *3*(6), 667–674. https://doi.org/10.1177/1948550611435941

Feingold, A. (1990). Gender differences in effects of physical attractiveness on romantic attraction: A comparison across five research paradigms. *Journal of Personality and Social Psychology*, *59*(5), 981–993. https://doi.org/10.1037/0022-3514.59.5.981

Feingold, A. (1992). Gender differences in mate selection preferences: A test of the parental investment model. *Psychological Bulletin*, *112*(1), 125–139. https://doi.org/10.1037/0033-2909.112.1.125

Field, A., Miles, J., & Field, Z. (2014). *Discovering Statistics Using R*. SAGE Publications Inc.

Fletcher, G. J. O., Kerr, P. S. G., 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*(4), 540–550. https://doi.org/10.1177/0146167213519481

Fletcher, G. J. O., Overall, Nickola, C., & Campbell, L. (2020). Reconsidering "Best Practices" for Testing the Ideal Standards Model: A Response to Eastwick, Finkel, and Simpson (2018). *Personality and Social Psychology Bulletin*, *46*(11), 1581–1595. https://doi.org/10.1177/0146167220910323

Fletcher, G. J. O., Simpson, J. A., Campbell, L., & Overall, N. C. (2019). *The science of intimate relationships* (2nd ed.). Wiley-Blackwell.

Fletcher, G. J. O., Simpson, J. A., & Thomas, G. (2000). Ideals, perceptions, and evaluations in early relationship development. *Journal of Personality and Social Psychology*, *79*(6), 933–940. https://doi.org/10.1037/0022-3514.79.6.933

Fletcher, G. J. O., Simpson, J. A., Thomas, G., & Giles, L. (1999). Ideals in intimate relationships. *Journal of Personality and Social Psychology*, *76*(1), 72–89. https://doi.org/10.1037/0022-3514.76.1.72

Frost, D. M., & Forrester, C. (2013). Closeness Discrepancies in Romantic Relationships: Implications for Relational Well-Being, Stability, and Mental Health. *Personality and Social Psychology Bulletin*, *39*(4), 456–469. https://doi.org/10.1177/0146167213476896

Funder, D. C. (2001). Three trends in current research on person perception: Positivity, realism, and sophistication. In J. A. Hall & F. J. Bernieri (Eds.), *Interpersonal sensitivity: Theory and measurement* (pp. 319–331). Psychology Press. https://doi.org/10.4324/9781410600424-27

Gerlach, T. M., Arslan, R. C., Schultze, T., Reinhard, S. K., & Penke, L. (2019). Predictive validity and adjustment of ideal partner preferences across the transition into romantic relationships. *Journal of Personality and Social Psychology*, *116*(2), 313–330. https://doi.org/10.1037/pspp0000170

Gerlach, T. M., Schultze, T., Arslan, R. C., & Penke, L. (2019, September 17). Do preference partner discrepancies prospectively predict the stability and quality of newly formed relationships? In T. M. Gerlach & L. Penke (Chairs), Personality and Romantic Outcomes. Symposium conducted at the meeting of 15th Meeting of the Section Differential/Personality Psychology and Psychological Assessment, Dresden, Germany.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2-3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Hofer, G., Burkart, R., Langmann, L., & Neubauer, A. C. (2021). What you see is what you want to get: Perceived abilities outperform objective test performance in predicting mate appeal in speed dating. *Journal of Research in Personality*, *93*, 104113. https://doi.org/10.1016/j.jrp.2021.104113

Holt-Lunstad, J., Birmingham, W., & Jones, B. Q. (2008). Is There Something Unique about Marriage? The Relative Impact of Marital Status, Relationship Quality, and Network Social Support on Ambulatory Blood Pressure and Mental Health. *Annals of Behavioral Medicine*, *35*(2), 239–244. https://doi.org/10.1007/s12160-008-9018-y

Joel, S., Eastwick, P. W., Allison, C. J., Arriaga, X. B., Baker, Z. G., Bar-Kalifa, E., Bergeron, S., Birnbaum, G. E., Brock, R. L., Brumbaugh, C. C., Carmichael, C. L., Chen, S., Clarke, J., Cobb, R. J., Coolsen, M. K., Davis, J., Jong, D. C. de, Debrot, A., DeHaas, E. C., . . . Wolf, S. (2020). Machine learning uncovers the most robust self-report predictors of relationship quality across 43 longitudinal couples studies. *Proceedings of the National Academy of Sciences*, *117*(32), 19061–19071. https://doi.org/10.1073/pnas.1917036117

Kurzban, R., & Weeden, J. (2005). HurryDate: Mate preferences in action. *Evolution and Human Behavior*, *26*(3), 227–244. https://doi.org/10.1016/j.evolhumbehav.2004.08.012

Kuznetsova, A., Brockhoff, P., & Christensen, R [R.] (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Lackenbauer, S. D., & Campbell, L. (2012). Measuring up: The unique emotional and regulatory outcomes of different perceived partner-ideal discrepancies in romantic relationships. *Journal of Personality and Social Psychology*, *103*(3), 472–488. https://doi.org/10.1037/a0029054

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*(6), 703–722. https://doi.org/10.1177/0146167216637842

Li, N. P., Bailey, J. M., Kenrick, D. T., & Linsenmeier, J. A. W. (2002). The necessities and luxuries of mate preferences: Testing the tradeoffs. *Journal of Personality and Social Psychology*, *82*(6), 947–955. https://doi.org/10.1037/0022-3514.82.6.947

Lüdecke, D. (2021). *sjPlot: Data Visualization for Statistics in Social Science* (Version 2.8.9.) [R package]. https://CRAN.R-project.org/package=sjPlot

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*(4), 933–964. https://doi.org/10.1111/j.1467-6494.2009.00570.x

Miller, G. F., & Todd, P. M. (1998). Mate choice turns cognitive. *Trends in Cognitive Sciences*, *2*(5), 190–198. https://doi.org/10.1016/S1364-6613(98)01169-3

Miller, R. B., Hollist, C. S., Olsen, J., & Law, D. (2013). Marital Quality and Health Over 20 Years: A Growth Curve Analysis. *Journal of Marriage and Family*, *75*(3), 667–680. https://doi.org/10.1111/jomf.12025

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*(4), 662–685. https://doi.org/10.1037/0022-3514.91.4.662

Park, Y., Impett, E. A., & MacDonald, G. (2020). Singles' Sexual Satisfaction is Associated With More Satisfaction With Singlehood and Less Interest in Marriage. *Personality and Social Psychology Bulletin*, *47*(5), 741–752. https://doi.org/10.1177/0146167220942361

Penke, L., & Asendorpf, J. B. (2008). Beyond global sociosexual orientations: A more differentiated look at sociosexuality and its effects on courtship and romantic relationships. *Journal of Personality and Social Psychology*, *95*(5), 1113–1135. https://doi.org/10.1037/0022-3514.95.5.1113

Pietromonaco, P. R., & Collins, N. L. (2017). Interpersonal mechanisms linking close relationships to health. *American Psychologist*, *72*(6), 531–542. https://doi.org/10.1037/amp0000129

R Core Team. (2020). *R: A language and environment for statistical programming* [Computer software]. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/

Rusbult, C. E., Arriaga, X. B., & Agnew, C. R. (2001). Interdependence in close relationships. In G. J. O. Fletcher & M. S. Clark (Eds.), *Blackwell Handbook of Social Psychology: Interpersonal Processes* (pp. 359–387). Blackwell Publishers.

Rusbult, C. E., Martz, J. M., & Agnew, C. R. (1998). The Investment Model Scale: Measuring commitment level, satisfaction level, quality of alternatives, and investment size. *Personal Relationships*, *5*(4), 357–387. https://doi.org/10.1111/j.1475-6811.1998.tb00177.x

Schwarz, S., & Hassebrauck, M. (2012). Sex and Age Differences in Mate-Selection Preferences. *Human Nature*, *23*(4), 447–466. https://doi.org/10.1007/s12110-012-9152-x

Therneau, T. M. (2020). *coxme: Mixed Effects Cox Models* (Version 2.2-16) [R package]. https://cran.rstudio.com/web/packages/coxme/index.html

Todd, P. M., & Miller, G. F. (1999). From pride and prejudice to persuasion: Satisficing in mate search. In G. Gigerenzer, P. M. Todd, & The ABC Research Group (Eds.), *Simple heuristics that make us smart* (pp. 287–308). Oxford University Press.

Todd, P. M., Penke, L., Fasolo, B., & Lenton, A. P. (2007). Different cognitive processes underlie human mate choices and mate preferences. *Proceedings of the National Academy of Sciences*, *104*(38), 15011–15016. https://doi.org/10.1073/pnas.0705290104

Tully, S., & Meyvis, T. (2017). Forgetting to remember our experiences: People overestimate how much they will retrospect about personal events. *Journal of Personality and Social Psychology*, *113*(6), 878–891. https://doi.org/10.1037/pspa0000094

Walter, K. V., Conroy-Beam, D., Buss, D. M., Asao, K., Sorokowska, A., Sorokowski, P., Aavik, T., Akello, G., Alhabahba, M. M., Alm, C., Amjad, N., Anjum, A., Atama, C. S., Atamtürk Duyar, D., Ayebare, R., Batres, C., Bendixen, M., Bensafia, A., Bizumic, B., . . . Zupančič, M. (2020). Sex Differences in Mate Preferences Across 45 Countries: A Large-Scale Replication. *Psychological Science*, *31*(4), 408–423. https://doi.org/10.1177/0956797620904154

Watson, D., Hubbard, B., & Wiese, D. (2000). Self–other agreement in personality and affectivity: The role of acquaintanceship, trait visibility, and assumed similarity. *Journal of Personality and Social Psychology*, *78*(3), 546–558. https://doi.org/10.1037/0022-3514.78.3.546

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org

Wieczorek, L. L., Tata, C. S., Penke, L., & Gerlach, T. M. (2020). Online implementation of an event history calendar with formr: A tutorial. *Personal Relationships*, *27*(1), 176–208. https://doi.org/10.1111/pere.12305

Wood, D., & Furr, R. M. (2016). The Correlates of Similarity Estimates Are Often
    Misleadingly Positive:The Nature and Scope of the Problem, and Some Solutions.
    Personality and Social Psychology Review, *20*(2), 79–99.
    https://doi.org/10.1177/1088868315581119

Supplement

Ideals are Associated with Perceived Traits of Future Partners Across 13 Years, With a Closer Match Being Related to Commitment

Julie C. Driebe[1], Julia Stern[1,3], Lars Penke[1,2], Tanja M. Gerlach[1,2]

[1]University of Goettingen

[2]Leibniz Science Campus Primate Cognition

[3]University of Bremen

S1. Deviations from our Pre-registration

Table S1

Description and Explanation of all Deviations to Our Pre-registration

Pre-registration Hypotheses	Manuscript	Explanation
We pre-registered the following hypotheses: ● 2.1. Initial ideal partner preferences (T1) correlate positively with characteristics of all future partners. ● 2.2. Initial ideal partner preferences (T1) are more strongly correlated with characteristics of partners with closer proximity to our initial assessment time point. ● 2.3. Current ideal partner preferences (T2) are more strongly correlated with characteristics of partners with closer proximity to our current assessment time point. ● 3.1 Relationship satisfaction is negatively associated with the discrepancy between ideal partner preferences and characteristics of a partner. ● 3.2. Relationship commitment is negatively associated with the discrepancy between ideal partner preferences and characteristics of a current partner. ● 3.3. Relationship dissolution being more initiated by the focal person is positively associated with the discrepancy between ideal partner preference and characteristics of the former partner. ● 3.4. Relationship duration is negatively associated with the discrepancy between ideal partner preferences and characteristics of the former partner.	We included all hypotheses as pre-registered in our manuscript except for 2.2 and 2.3.	We do not think that our data allow us to answer 2.2 and 2.3. Note that hypotheses are numbered as they were in the pre-registration but are numbered differently in the manuscript.

| Pre-registration	Manuscript	Explanation
Analyses		
We preregistered that our final sample would consist out of both initial samples.	We analysed both samples separately, whereas we only report analyses based on the former BSDS in our main manuscript. Analyses based on the former sample of the Sociosexuality Study are reported in our supplement (see S3).	Assessment of T1 ideal partner preferences differed tremendously between the two initial samples. This made it impossible to combine the two in our analyses. For example the range of the Euclidean distance differs between the two as the number of items influence the measure. We decided to only report results of the sample of former participants of the Sociosexuality Study online because the sample is too small, though limited in power to draw reliable conclusions. This is why we only report results based on this sample online for the sake of completeness.
For hypothesis 2.1, we pre-registered to specify a random effect in which partners (level 1) are nested within each participant (level 2) which are nested within the study of their initial participation (level 3).	We specified a random intercept for partners, participants and items.	The diverging assessment of T1 ideal partner preferences has led to different variances between the two studies. Including a random intercept for the initial study is therefore not appropriate. Instead the two initial samples were analysed separately.
We preregistered to run additional analyses investigating ideal partner preference-matching.	We did not include analyses based on the level-metric.	When calculating the level-metric, a relationship outcome is predicted with participant's ideal partner preference and their perceived partner characteristic. This is mostly calculated for each preference dimension. Although not part of the current manuscript, we determined 8 preference dimensions. This means that we would need to calculate 72 models (9 relationship outcomes x 8

|		
		preference dimensions). However, we believe that our calculated preference dimensions have a large measurement error since they were determined in a different sample and were assessed with a long-version assessing participant's ideal partner preferences. We believe that this measurement error is particularly severe in the large number of models we would have needed to calculate, making these analyses prone to false-positive results.
We preregistered to calculate the raw and corrected pattern metric.	We calculated the corrected pattern metric but not the raw pattern metric. Instead we calculated a partly corrected pattern metric. These analyses calculate the within-person correlation between ideals and partner perceptions. Within-person correlations are then Fisher z-transformed and used to predict relationship outcomes. We did not Fisher z-transform correlation coefficients.	We only analysed the corrected pattern metric as it has been proposed to be more appropriate compared to the raw pattern metric (Eastwick et al., 2019). Correlation coefficients are Fisher z-transformed for a more normal distribution. However, a normal distribution is not necessary plus correlation coefficients were already normally distributed.
Relationship Length	We analysed the relationship length using a survival model.	Using a survival model enabled us to include more data increasing our power and is thus preferable.

S2. Further Descriptives

Figure S1

Participant Flow
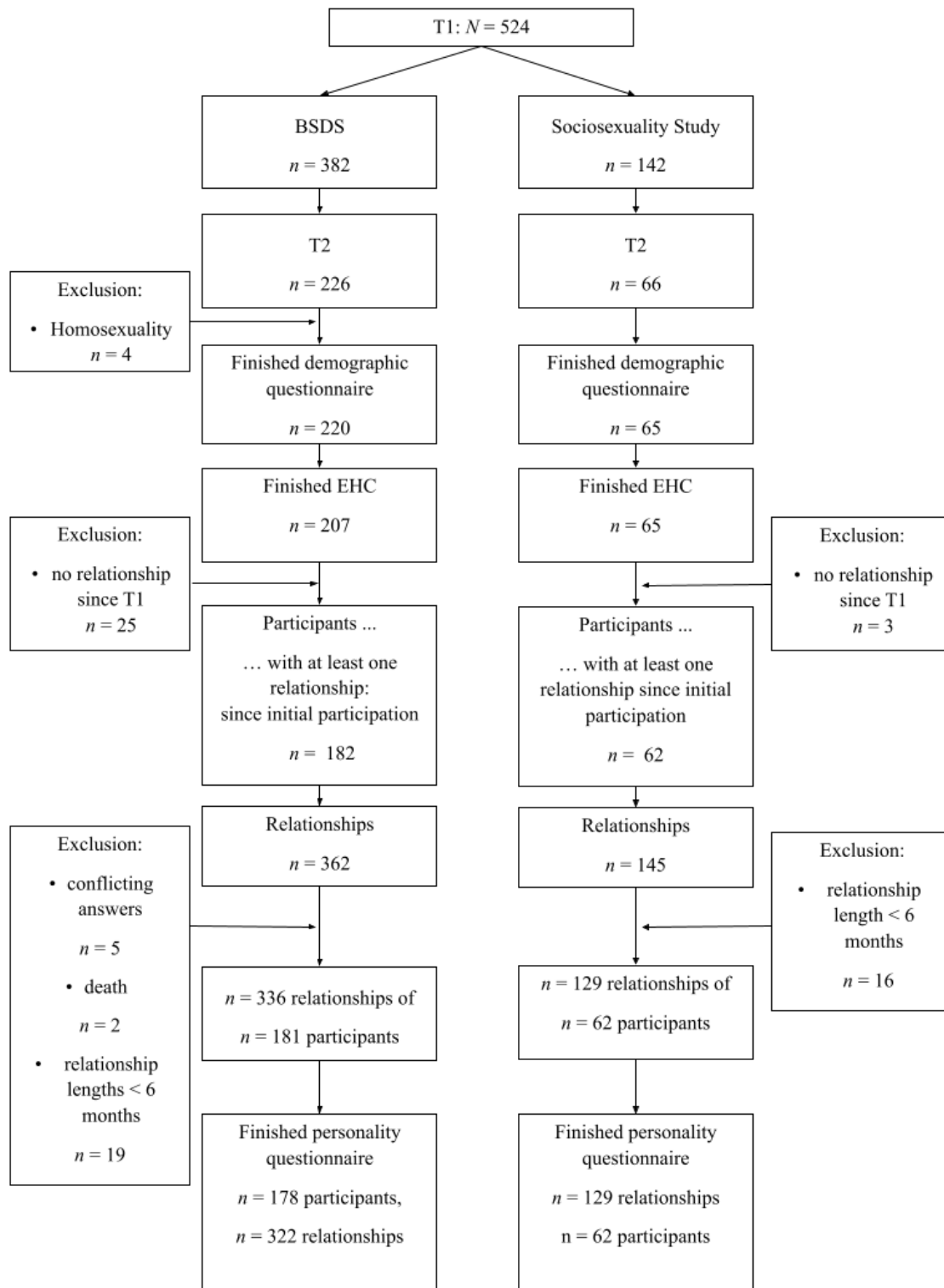
Table S2

Attrition analysis

Variable	Dropout n = 156 M (SD)	Participants n = 226 M (SD)	p-value	Hedges' g
Age (years)	33.06 (7.80)	32.56 (7.05)	=.526	-.07
Openness	3.82 (.53)	3.83 (.50)	=.873	.02
Conscientiousness	**3.89 (.60)**	**3.77 (.62)**	**=.049**	**-.20**
Extraversion	3.57 (.52)	3.47 (.60)	=.108	-.16
Agreeableness	3.81 (.50)	3.81 (.54)	=.990	.00
Neuroticism	**2.52 (.70)**	**2.69 (.72)**	**=.026**	**.23**

Note. Using two-tailed t-tests, we compared T1 values of participants who still participated in our study at T2 with participants who dropped out from T1 to T2.

Figure S2
Correlation coefficients between T1 ideal partner preferences and perceived partner traits
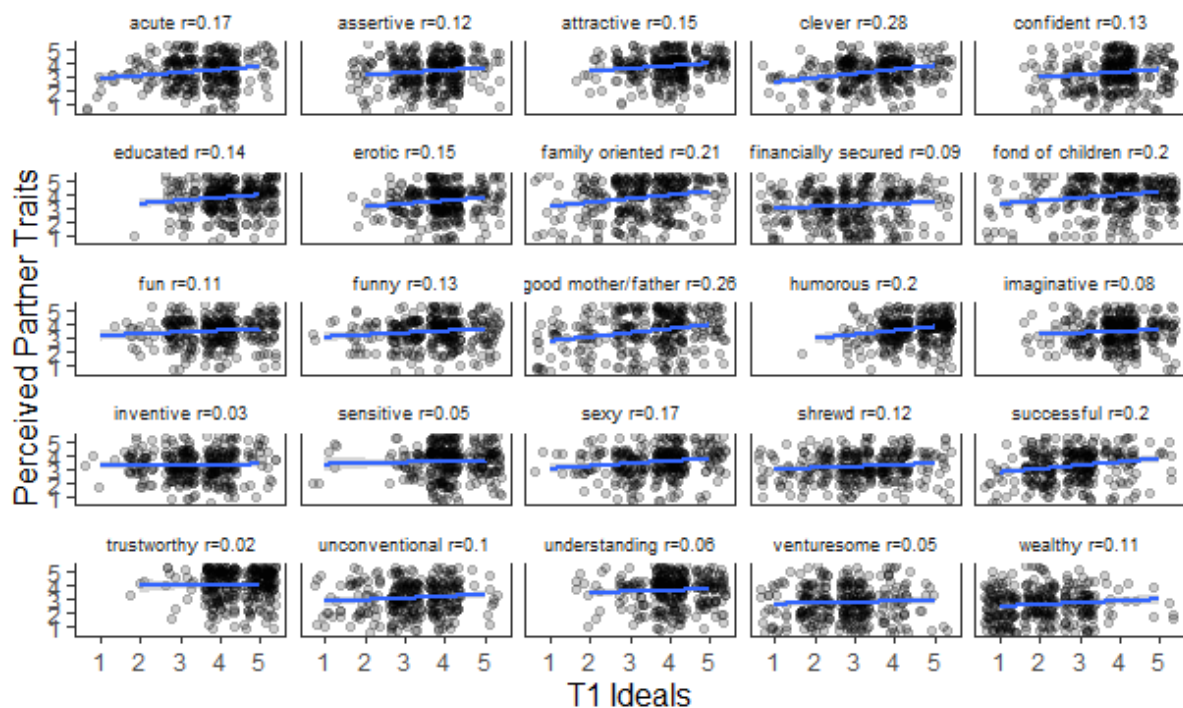
Note. These correlation coefficients need to be interpreted with caution as they ignore the hierarchical structure of our data. The figure demonstrates that the correlation between ideals and perceived partner characteristics differed between items. The solid line illustrates the regression line with the shaded area in grey representing the 95% confidence interval. The dots represent the participant's estimates whereas varying opacity is caused by multiple participants overlapping.

S3. Sociosexuality Study

Method

Sample

From 2004 to 2005 (T1), 284 individuals participated in the study entitled *Love, Sexuality and Personality*, herein referred to as Sociosexuality Study (Penke & Asendorpf, 2008). Of these initial participants, only 142 single participants were asked to report their ideal partner preferences which we invited to participate in our online study (T2). We were able to recruit 66 participants (dropout 54%). Because we had to exclude participants who did not finish our study ($n = 1$) and reported not to have a relationship exceeding 6 months since T1 or were not currently involved in a romantic relationship ($n = 3$), our final sample consists out of 62 participants ($n = 30$ female) with a mean age of 38.7 years ($SD = 3.3$, range = 34 - 45 years). Most (76%) participants are currently involved in a relationship. The highest level of education was for the majority a university degree (81%) and for 19% a school degree. Participants rated 129 partners (69 female) of which 65% were with former partners and 35% with current partners. Partners were on average 39.2 years old (SD = 7.7, range = 25 - 80 years).

Procedure and Measures

Because participants were invited to the same online study, the study procedure was exactly the same as described for the BSDS in our main manuscript. The only difference between the samples is the assessment of T1 ideal partner preferences. Instead of assessing the importance of each characteristic, participants of the Sociosexuality Study rated the desired level of 13 characteristics of an ideal partner. Thus, participants indicated how strongly they would like a certain characteristic to be expressed in an ideal partner on a scale from 1 (very little) to 5 (as strong as possible). At T2, participants rated each partner they listed in the EHC whether he or she possessed several characteristics. 7 of these attributes corresponded to the ideal partner preferences assessed at T1.

Table S3

Means and Standard Deviations of Perceived Partner Traits, T1 Ideal Partner Preferences and Relationship Outcomes.

Item	Traits	T1 Ideals
	M (*SD*)	*M* (*SD*)
occupational success, good occupational perspective	3.15 (1.28)	3.26 (0.65)
faithfulness, reliability	3.86 (1.30)	3.50 (0.78)
body attractiveness, sex-appeal	3.77 (0.89)	3.58 (0.69)
parental abilities, fondness of children	3.64 (1.22)	3.15 (0.81)
intelligence	3.88 (1.02)	3.35 (0.73)
creativity	3.38 (1.03)	3.58 (0.80)
thirst for adventure, activeness	3.47 (1.08)	3.58 (0.64)
Relationship Outcomes	M (SD)	
general satisfaction	3.56 (0.92)	
harmony	3.61 (1.03)	
family	3.40 (1.56)	
passion	3.36 (1.21)	
standard of living	3.65 (1.08)	
reliability	4.03 (1.19)	
Commitment (Cronbach's α = 0.94)	4.34 (1.02)	
Relationship Length	34.80 (36.03)	
Break-up	3.38 (1.61)	

Note. Mean commitment is based on current relationships only ($n = 45$) and mean relationship lengths and mean break-up is based on former relationships ($n = 84$).

Results

We summarise results found in the sample of former participants of the Sociosexuality Study, but do not interpret these due to the small sample size. Full tables can be found on the OSF (https://osf.io/tyc4r/?view_only=c0ea03dbbb004814b92426c3a31b9b3a).

Ideal-Trait Correlation

We found a positive association between perceived partner characteristics and ideal partner preferences reported at T1 (Table S4).

Table S4

Multilevel Model Predicting Perceived Partner Characteristics with Participant's T1 Ideals, Controlling for Age and Sex

Coeffcient	Perceived Partner Characteristics				
	Estimates	std. Error	95% CI	Statistic	p-values
Intercept	3.53	0.13	3.28 – 3.79	27.31	<0.001
T1 Ideals (z-standardised)	0.12	0.04	0.05 – 0.20	3.26	0.001
Age (z-standardised)	-0.16	0.06	-0.28 – -0.05	-2.80	0.005
Gender (0 = women, 1 = men)	0.19	0.12	-0.05 – 0.42	1.57	0.116
Random Effects					
σ^2			1.04		
$\tau_{00 \; partner\_name\_unique}$			0.10		
$\tau_{00 \; session}$			0.07		
$\tau_{00 \; item}$			0.07		
ICC			0.19		
$N_{session}$			62		
$N_{partner\_name\_unique}$			129		
N_{item}			7		
Observations			903		
Marginal R^2 / Conditional R^2			0.035 / 0.215		

Ideal Partner Preference-Matching

(Semi-) Euclidean Distance

For the Euclidean distance, only one relationship outcome was significant, namely whether participants relied on their (former) partners. This was the only outcome which showed significant effects across all analyses (Table S5).

Regarding the Semi-Euclidean distance, we found significant effects for five relationship outcomes, of which all but one remained significant when including a score of how positive participants rated their partners.

Table S5

Estimates and Confidence Intervals of the (Semi)-Euclidean Distance Predicting Each Relationship Outcome

Outcome	Estimate of the Euclidean distance [95% CI]		Estimate of the Semi-Euclidean distance [95% CI]	
	Main Model	with Positivity	Main Model	with Positivity
Satisfaction	0.07 [-0.08 – 0.23]	0.07 [-0.07 – 0.22]	0.29*** [0.14 – 0.44]	0.14 [-0.04 – 0.31]
Harmony	-0.00 [-0.18 – 0.18]	0.00 [-0.17 – 0.17]	0.14 [-0.04 – 0.32]	0.00 [-0.17 – 0.17]
Family	0.10 [-0.17 – 0.37]	0.10 [-0.16 – 0.36]	0.53*** [-0.16 – 0.36]	0.35* [0.03 – 0.66]
Passion	-0.06 [-0.27 – 0.15]	-0.06 [-0.27 – 0.15]	0.03 [-0.15 – 0.21]	-0.14 [-0.40 – 0.11]
Living	-0.01 [-0.19 – 0.17]	0.00 [-0.17 – 0.18]	0.36*** [0.19 – 0.53]	0.25* [0.04 – 0.46]
Reliability	0.26* [0.06 – 0.47]	0.26** [0.07 – 0.46]	0.49*** [0.29 – 0.69]	0.45*** [0.21 – 0.69]
Commitment	-0.19 [-0.58 – 0.19]	-0.15 [-0.54 – 0.24]	0.15 [-0.24 – 0.55]	0.01 [-0.45 – 0.47]
Relationship Length	-0.04 [-0.29 – 0.21]	-0.04 [-0.29 – 0.20]	-0.54*** [-0.78 – -0.30]	-0.35* [-0.65 – -0.06]
Break-up	0.06 [-0.26 – 0.37]	0.06 [-0.26 – 0.37]	-0.01 [-0.35 – 0.33]	-0.03 [-0.35 – 0.33]

Note. *** p <. 001, ** p <.01 , * p < .05. [95% CI] = 95% confidence intervals. Distance scores were *z*-standardised before predicting each relationship outcome. In the column "with positivity", the estimate of the (semi)-Euclidean distance on each relationship outcome is listed while controlling the model on how positive each partner was rated. Full models can be found at: https://osf.io/tyc4r/?view_only=c0ea03dbbb004814b92426c3a31b9b3a.

(Partly) Corrected Pattern Metric

When predicting each relationship outcome with the within-person correlation we found no significant effects. The only exception is the outcome in which participants reported their satisfaction with their standard of living with a (former) partner when predicted with the partly corrected pattern metric (norm-centring perceived partner characteristics), but not the

corrected pattern metric (norm-centring perceived partner characteristics and ideal partner preferences).

Table S6

Estimates of the (Partly) Corrected Pattern Metric Predicting Each Relationship Outcome

Outcome	partly corrected pattern metric	corrected pattern metric
Satisfaction	0.07 [-0.34 – 0.47]	0.13 [-0.28 – 0.54]
Harmony	-0.20 [-0.65 – 0.25]	-0.17 [-0.63 – 0.29]
Family	-0.28 [-0.96 – 0.41]	0.07 [-0.62 – 0.77]
Passion	0.28 [-0.25 – 0.81]	0.26 [-0.28 – 0.80]
Living	-0.46* [-0.89 – -0.03]	-0.24 [-0.69 – 0.21]
Reliability	-0.02 [-0.54 – 0.50]	0.17 [-0.36 – 0.70]
Commitment	0.01 [-0.80 – 0.82]	0.01 [-0.82 – 0.84]
Relationship Length	0.50 [-0.10 – 1.09]	0.21 [-0.78 – 0.36]
Break-up	-0.25 [-1.08 – 0.58]	-0.50 [-0.37 – 0.79]

Note. *** p <. 001, ** p <.01 , * p < .05. [95% CI] = 95% confidence intervals. Full models can be found at: https://osf.io/tyc4r/?view_only=c0ea03dbbb004814b92426c3a31b9b3a

Appendix D.

Curriculum Vitae

[Curriculum vitae omitted from online version]