Evolution and Human Behavior 42 (2021) 507–516

Contents lists available at ScienceDirect

Evolution and Human Behavior

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

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

Julie C. Driebe [a,1,*], Morgan J. Sidari [b,1,*], Michael Dufner [c], Juliane M. von der Heiden [d], Paul C. Bürkner [e], Lars Penke [a,f,2], Brendan P. Zietsch [b,2], Ruben C. Arslan [g,2]

[a] Biological Personality Psychology, University of Goettingen, Gosslerstrasse 14, 37073 Goettingen, Germany
[b] Centre for Psychology and Evolution, McElwain Building (24A), School of Psychology, University of Queensland, St Lucia 4072, Australia
[c] Personality Psychology and Psychological Assessment, Witten/Herdecke University, Alfred-Herrhausen-Straße 50, 58448 Witten, Germany
[d] Personality Psychology and Psychological Assessment, University of Mainz, Bingener Str. 14 – 16, 55099 Mainz, Germany
[e] Computer Science, Aalto University, Konemiehentie 2, 02150 Espoo, Finland
[f] Leibniz Science Campus Primate Cognition, Goettingen, Deutsches Primatenzentrum GmbH, Kellnerweg 4, 37077 Göttingen, Germany
[g] Center for Adaptive Rationality, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany

ARTICLE INFO

Keywords:
Intelligence
Mate choice
Sexual selection

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 ($N = 763$), 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.

1. Introduction

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, Blumenschine, & Adams, 1981). One theory is that our surplus of intelligence has emerged through intersexual selection (Miller, 2000a; Miller, 2000b), 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, 2000a; Miller, 2000b). 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, 2000a; Miller, 2000b; 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, Bailey, Kenrick, & Linsenmeier, 2002; Sprecher & Regan, 2002; Wilbur & Campbell, 2011); however, other studies have shown that these ideal partner preferences do not closely correspond to mate preferences

* Corresponding authors.
E-mail addresses: julie.driebe@uni-goettingen.de (J.C. Driebe), morgan.sidari@uqconnect.edu.au (M.J. Sidari).
[1] Share first authorship.
[2] Share last authorship.

https://doi.org/10.1016/j.evolhumbehav.2021.05.002
Received 1 December 2020; Received in revised form 9 April 2021; Accepted 12 May 2021
Available online 31 May 2021

revealed in attraction to real individuals (Eastwick, Eagly, Finkel, & Johnson, 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, Coss, Scheib, and Blozis (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.

2. The current study

Here, we conducted two studies to investigate the accuracy of intelligence judgments 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 judgments 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.

2.1. 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.

Overview of study 1 stimuli.

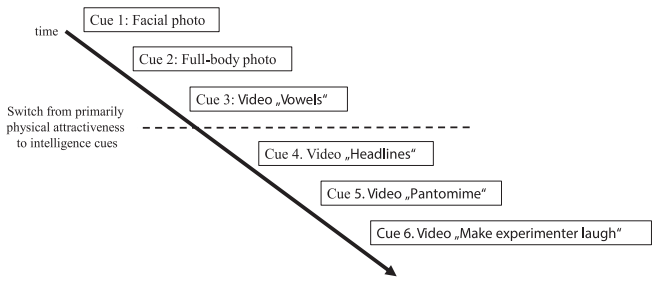

Fig. 1. Overview of study 1 stimuli.

2.2. 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.

3. Method

3.1. Study 1

Parts of study 1 were preregistered (https://osf.io/q7sw5/); 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.

3.1.1. Participants

3.1.1.1. 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 Goettingen university campus. Of the 347 participants that commenced the survey, 118 males finished.[3] 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 school leaving certificates (2%). The vast majority of the sample was heterosexual (97%), with one homosexual and two bisexual participants. The

[3] 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.

majority of men were single (61%) and the remainder were currently in a romantic relationship (39%).

3.1.1.2. 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).

3.1.2. Materials and procedure

3.1.2.1. Male targets. Participants completed an online questionnaire implemented via the survey framework formr.org (Arslan, Tata, & Walther, 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, Liewald, & Nissan, 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, Mauer, Riemann, Spinath, & Angleitner, 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 s 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 × 1080 pixels) were used for all recordings and clips were created with the program Mangold VideoSyncPro IP Version 1.7.0.22.

3.1.2.2. 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).

3.1.2.3. 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).

3.1.2.4. 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).

3.1.2.5. 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).

3.1.2.6. 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).

3.1.2.7. 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.

3.1.2.8. 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).

3.1.3. Statistical analyses

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

3.1.3.1. 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.

3.1.3.2. 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.

3.1.3.3. 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, Mächler, Bolker, & Walker, 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.

3.1.3.4. 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, Graham, & Zeileis, 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.

3.1.3.5. 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.

3.2. Study 2

3.2.1. Participants

Participants were 763 (397 female) first year psychology students with ages ranging from 16.92 to 38.91 years (women: $M = 19.27$, $SD = 2.67$; men: $M = 19.79$, $SD = 2.63$). Participants were recruited between 2016 and 2019 from the University of Queensland'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 (4.85%) or in a committed relationship (6.17%) 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 45.49% of the time and they mutually said yes 21.05% 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 Queensland (Ethics #16-PSYCH-4-65-JS).

3.2.2. 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.

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

3.2.2.2. 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*.

3.2.2.3. 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'.

3.2.3. Procedure

3.2.3.1. 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 1.7 m 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.

3.2.3.2. 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 125 speed-dating sessions with 763 participants. Participants participated in 2–4 dates ($M = 3.08$).

3.2.3.3. 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.

3.2.4. 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 modelling (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.

4. Results

4.1. Study 1

4.1.1. 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.

4.1.2. 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, $r = 0.34$, ($p < .001$; 95% CI [0.14; 0.51], 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 ($\beta = 0.22$, $p < .001$, 95% CI [0.07; 0.28]) 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).

4.1.3. 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 ($r = 0.92$). 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).

4.1.4. 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 ($b = 0.30$, $p > .001$, 95% CI [0.24; 0.36]). But contrary to expectations, measured intelligence was not associated with perceived funniness ($r = -0.14$, $p = .18$, 95% CI [−0.34; 0.07], 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 ($b = -0.14$, $p = .03$, 95% CI [−0.26; −0.01]), contrary to expectations. However, men who were perceived to be funnier had a higher sexual mate appeal ($b = 0.35$, $p < .001$, 95% CI [0.26; 0.45]). 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 ($b = 0.17$, $p = .002$, 95% CI [0.06, 0.29]) (Table S11).

4.1.5. 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: $b = -0.07$, 95% HDI [− 0.11, −0.03]). 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.

4.1.6. 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: $b = 0.01$, 95% HDI [−0.02; 0.04]). This finding does not support our prediction, in that cues of intelligence did not uniquely contribute to sexual mate appeal ratings.

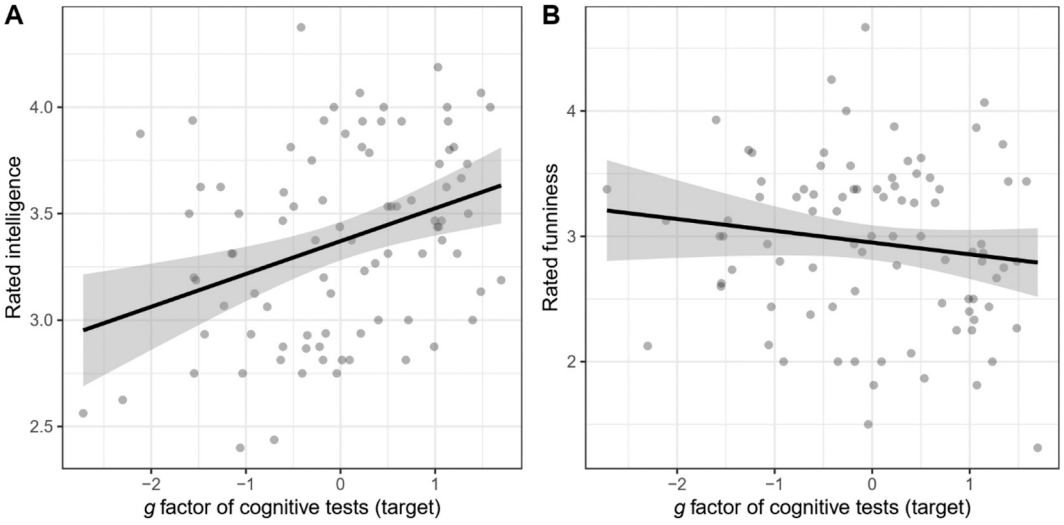

Fig. 2. Aggregated perceiver accuracy for intelligence as measured by the g factor. Note. The shaded area in grey reflects the 95% HDI.

Table 1
LM coefficients for associations between measured intelligence, humour and sexual mate appeal.

Term	Sexual mate appeal		
	Estimate	*p*	95% CI
Intercept	0.69	<0.001	[0.35; 1.03]
g factor	−0.14	0.03	[−0.26; −0.01]
Funniness	0.35	<0.001	[0.26; 0.45]
Physical attractiveness	0.24	<0.001	[0.17; 0.30]

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.

Table 2
Associations between sexual mate appeal and measured intelligence in sequential cue presentation.

Term	Sexual mate appeal	
	Estimate	95% HDI
Intercept	0.44	[0.09; 0.80]
Cue 1&2	−0.30	[−0.39; −0.21]
Cue 4	0.16	[0.09; 0.23]
Cue 6	0.36	[0.25; 0.48]
Physical attractiveness	1.15	[1.07; 1.22]
g factor	−0.07	[−0.10; −0.03]
Cue 1&2 * physical attractiveness	−0.10	[−0.13; −0.07]
Cue 4 * physical attractiveness	0.07	[0.04; 0.10]
Cue 6 * physical attractiveness	0.14	[0.11; 0.18]
Cue 1 & 2 * *g* Factor	−0.01	[−0.03; 0.02]
Cue 4 * *g* Factor	0.01	[−0.02; 0.04]
Cue 6 * *g* Factor	0.02	[−0.01; 0.05]

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, 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: $b = 0.02$, 95% HDI [−0.02; 0.04]). Taken together with the previous finding, this casts further doubt on the notion that intelligence is attractive in men.

4.1.7. 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: $b = 0.16$, 95% HDI [0.07; 0.24]) and further after cue 6 was presented (cue 6: $b = 0.36$, 95% HDI [0.23; 0.50]). 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: $b = 0.07$, 95% HDI [0.04; 0.11]) and the presentation of cue 6 (cue 6 x physical attractiveness: $b = 0.14$, 95% HDI [0.10; 0.19]). 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.

4.2. 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 [0.03; 0.13], $p = .002$). After aggregating ratings across raters, the correlation was $r = 0.12$ (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 [−0.06; 0.04], $p = .724$) (Fig. 5). We found no evidence that the associations between intelligence and perceptions differed by sex (*ps* > 0.88).

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. 6). We found no evidence that the associations with mate appeal differed by sex (*ps* > 0.38). 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.

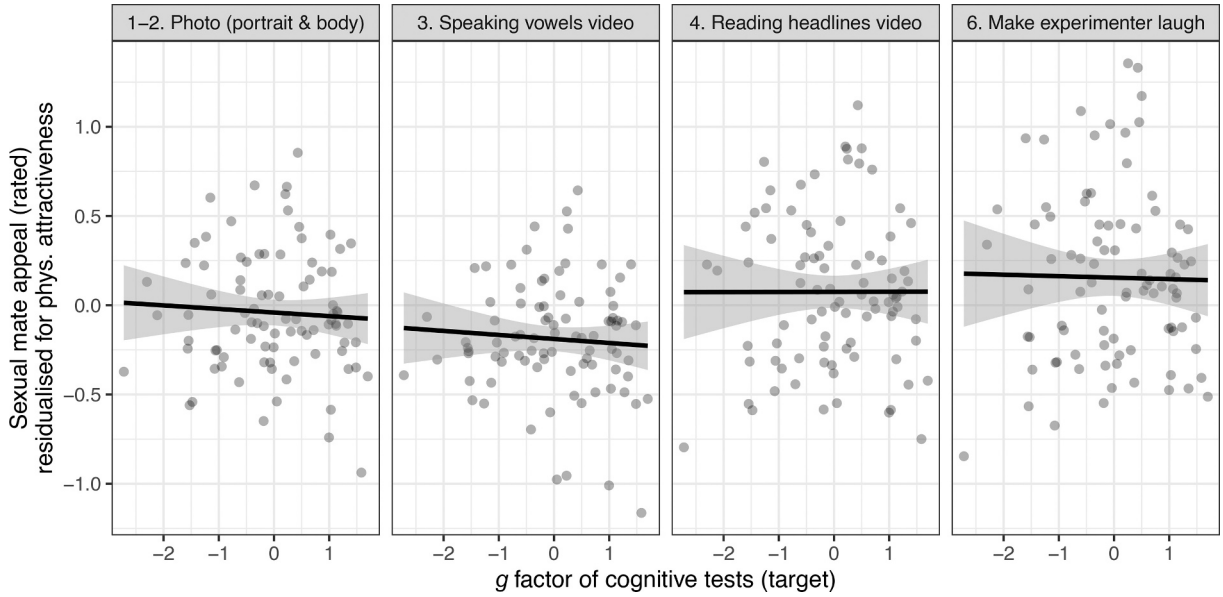

Fig. 3. The aggregated sexual mate appeal ratings made after seeing each cue (or set of cues) was adjusted for physical attractiveness. *Note.* The points shown in this plot show sexual mate appeal residualised for physical attractiveness. The shaded area in grey reflect the 95% HDI. The plot shows the 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.

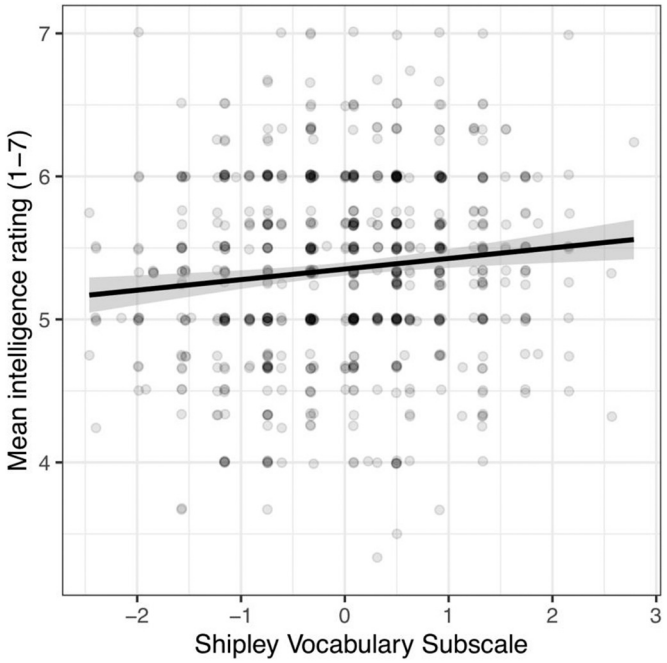

Fig. 4. Association between intelligence, as measured by the Shipley Institute of Living Scale (Vocabulary Subscale), and rated intelligence, after aggregating across raters. *Note.* Varying opacity of the dots is caused by overlap of multiple participants.

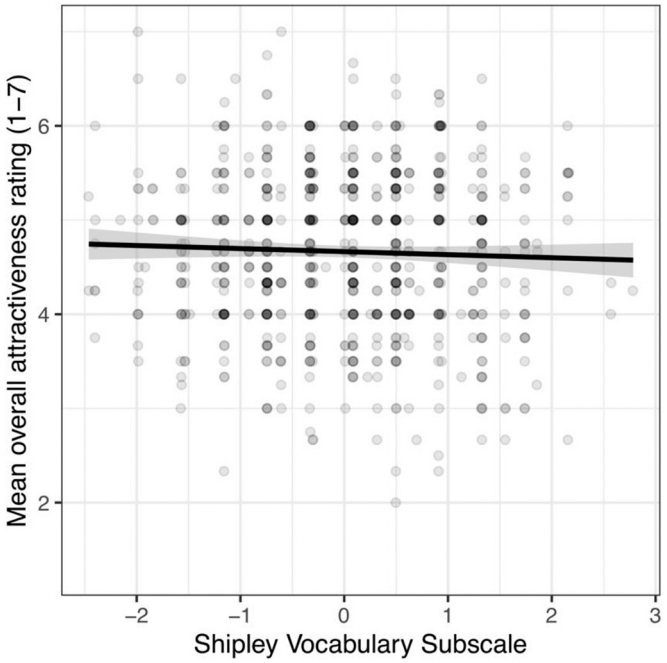

Fig. 5. The association between intelligence, as measured by the Shipley Institute of Living Scale (Vocabulary Subscale), and rated mate appeal, after aggregating across raters.

5. 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, 2000a; Miller, 2000b). 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 = 0.23) but not men's (r = 0.05) verbal intelligence with a factor representing mate quantity, while in the UK Biobank (Neale Lab UKBB, 2018) there is a negative genetic correlation of men's fluid intelligence and their number of sexual partners ($r = -0.18$, $p < .001$) but no significant genetic correlation for women (r = 0.07, 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

Table 3
MLM coefficients for associations between the Shipley Institute of Living Scale (Vocabulary Subscale), rated intelligence, rated funniness, and rated mate appeal.

Predictors	Rated Mate Appeal (1–7)				
	Estimates	*CI*	*p*	$N_{interactions}$	$N_{participants}$
Shipley (Vocabulary Subscale)	−0.01	−0.07–0.04	0.619	2245	753
Rated Intelligence	0.30	0.26–0.34	<0.001	2319	753
Rated Funniness	0.41	0.38–0.45	<0.001	2319	753

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

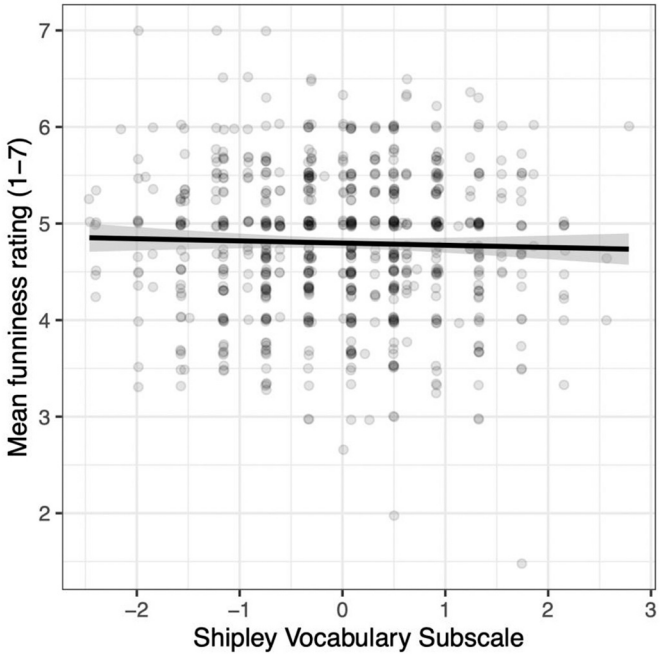

Fig. 6. The association between intelligence, as measured by the Shipley Institute of Living Scale (Vocabulary Subscale), and funniness, after aggregating across raters.

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, Hall, & Randal, 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, Hall, & LeBeau, 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.

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/q7sw5/?view_only=7e3293ddc0dd469aa7887d1e2fd64469). All analysis scripts have been made publicly available at the Open Science Framework (https://osf.io/rs3tg/?view_only=191205b275a2457086e2bf52727341d4).

Funding

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

Acknowledgements

We thank Julia Specht, Christina Hillebrand, Philipp Grochowski and Pia Drewke for help collecting the data of study 1. We thank Peter Zezula for his support setting up parts of study 1.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.evolhumbehav.2021.05.002.

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 J. A. Hall, S. M. Marianne, & T. V. West (Eds.), *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. 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.

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). https://doi.org/10.1177/147470490800600411, 147470490800600.

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.), *Evolutionary perspectives on human reproductive behavior: Vol. 907. Annals of the New York academy of sciences* (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. (2018). UKBB Genetic Correlation. n.d. http://www.nealelab.is/uk-biobank/

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 and 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 kristallinger 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 and 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.