Scientific Reports

https://doi.org/10.1038/s41598-026-52308-x

Article in Press

Large language models can detect verbal indicators of romantic attraction

Received: 3 April 2026

Accepted: 4 May 2026

Published online: 11 May 2026

Cite this article as: Matz S.C., Peters H., Cerf M. *et al.* Large language models can detect verbal indicators of romantic attraction. *Sci Rep* (2026). https://doi.org/10.1038/s41598-026-52308-x

Sandra C. Matz, Heinrich Peters, Moran Cerf, Eric Grunenberg, Paul W. Eastwick, Mitja Back & Eli J. Finkel

Title: Large Language Models Can Detect Verbal Indicators of Romantic Attraction

Authors:
Matz*, Sandra C.; Columbia Business School, United States, New York. (ORCID: 0000-0002-0969-4403)

Peters, Heinrich: Columbia Business School, United States, New York (ORCID: 0000-0002-0571-6388)

Cerf, Moran; Columbia Business School, United States, New York (ORCID: 0000-0002-2012-3177)

Grunenberg, Eric; Muenster University, Germany, Muenster (ORCID: 0000-0001-6296-468X)

Eastwick, Paul W.; University of California, Davis, United States, Davis (ORCID: 0000-0001-8512-8721)

Back, Mitja; Muenster University, Germany, Muenster (ORCID: 0000-0003-2186-1558)

Finkel, Eli J.: Northwestern University, United States, Evanston (ORCID: 0000-0002-0213-5318)

*Corresponding author: Sandra C. Matz (sm4409@gsb.columbia.edu)

Declarations:

- No funding was received for this study.
- The authors declare no conflict of interest.
- The project has received ethical approval from the Ethics Committee of Northwestern University (Protocol #1343-019)
- The manuscript was prepared without the assistance of AI tools.
- Data and analyses codes are available on OSF
- The manuscript is original, not previously published, and not under concurrent consideration.

Abstract

What makes people “click” on a first date and become mutually attracted to one another? While understanding and predicting the dynamics of romantic interactions used to be exclusive to human judgment, we show that Large Language Models (LLMs) demonstrate some capacity to detect

linguistic cues of romantic attraction during brief getting-to-know-you interactions. Examining data from 964 speed dates, we show that ChatGPT (and Claude 3) can predict both objective and subjective indicators of speed dating success (r=0.12-0.23), and that their judgements overlap with those made by human observers (r=0.21-0.35) with modest levels of accuracy. Notably, however, ChatGPT's predictions of actual matching (i.e., the exchange of contact information) were not only on par with those of human judges who had access to the same information but also incremental to speed daters' own predictions. Drawing on the Brunswik lens model, our findings also offer insights into how ChatGPT arrives at its judgements. Specifically, they suggest that its predictions can be explained by a combination of common content dimensions (e.g. the valence of the conversation) as well as more complex conversational dynamics (e.g., the use of humor, common interests or aligned values). While we found substantial overlap in the social cues utilized by ChatGPT and human raters, not all of these cues were valid predictors of matching. This suggests that both humans and LLMs rely on shared but imperfect heuristics when judging romantic attraction.

Keywords: Large Language Models, ChatGPT, speed-dating, romantic attraction, relational agents.

Can Large Language Models Detect Verbal Indicators of Romantic Attraction?

The dynamics of human interaction have long been a subject of interest for researchers seeking to understand romantic attraction[1]. Can we predict

whether two people are going to say “yes” to each other after an initial encounter? And what factors predict the elusive "clicking" that sparks interest between two people and prompts the possibility of a partnership?

The ability to analyze and interpret interpersonal dynamics has historically been confined to human perception, with research showing that observers can predict romantic interest[2,3] and long-term relationship success[4] with at least some level of accuracy. However, rapid advances in computers’ capacity to represent unstructured data – including text and (moving) images –have made them viable contestants with respect to understanding conversational dynamics and predicting the success of initial romantic encounters[5,6].

The most notable of these technological advances is the emergence of Large Language Models (LLMs[7–11]). LLMs are trained on vast corpora of textual data to learn statistical patterns in language[7] and produce novel text that is often indistinguishable from that created by humans[8]. By using probabilistic estimates for which words (or groups of words) are most likely to follow a particular prompt, LLMs can answer open-ended questions, summarize content, or translate text from one language to another.

However, LLMs are capable of far more than merely generating text. As a growing body of research suggests, LLMs possess human-like abilities that allow them to engage in complex social reasoning and interpret the types of social cues needed to understand human interactions beyond what is expressed on the surface. For example, LLMs exhibit properties resembling

theory of mind[9] the ability to accurately impute the mental states of other entities, and have demonstrated a remarkable ability to detect human emotion and sentiment[10–13] as well as psychological dispositions such as personality traits[14,15]. Similarly, research suggests that LLMs can discern subtle social signals, such as humor[16] or sarcasm[17], and both identify and explain social norms and norm violations in text, reflecting a capacity to understand implicit social rules[18].

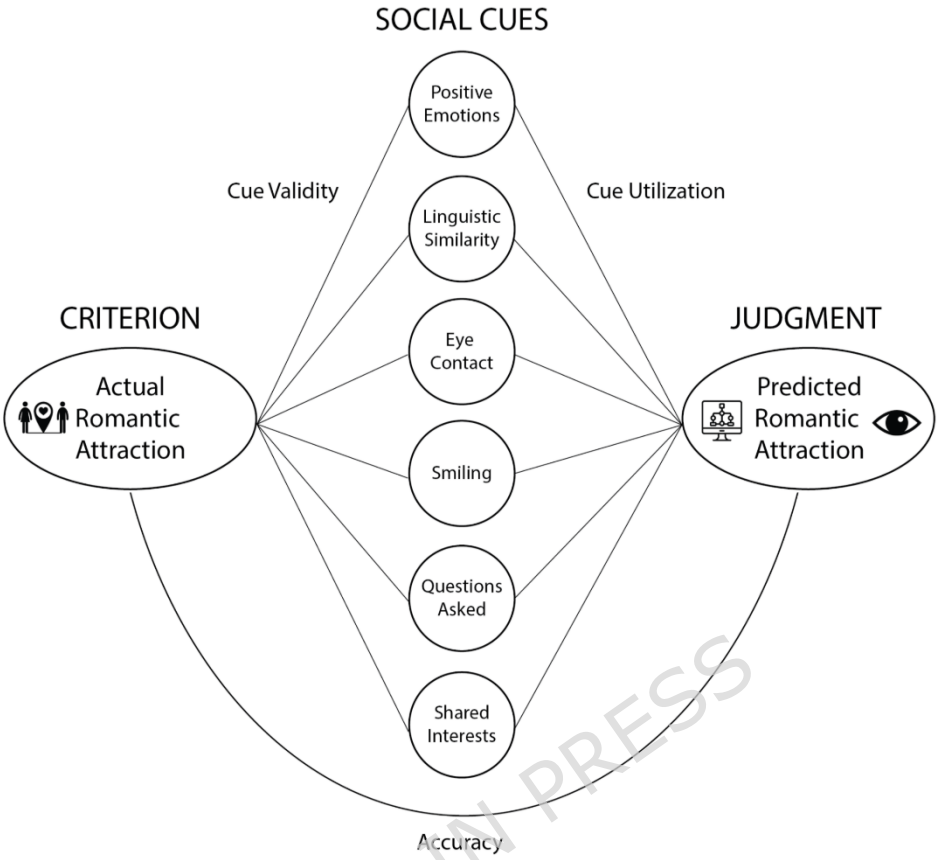
In this paper, we explore whether Large Language Models (LLMs) can detect external signals of romantic attraction during brief getting-to-know-you interactions between human subjects (i.e., speed dates). The speed-dating context offers an interesting test case for evaluating LLMs' capacity for social judgment for multiple reasons. First, speed dates are highly structured and time-limited in format, allowing for a direct comparison across multiple dyads. Second, the brevity of interactions makes our test a conservative estimate of LLMs' capacity to interpret the content of social interactions. Finally, the dynamic nature of the conversations allows LLMs to consider not just individual social cues but also their temporal sequence in the broader context of socio-cultural expectations and norms.

Our investigation is exploratory – rather than hypothesis-driven – in nature and follows a predictive – rather than explanatory – approach that is aimed at studying LLMs' baseline capacity for social judgements[19]. However, we ground this predictive exploration in the Brunswik's Lens model[20], a

conceptual framework for understanding and systematically analyzing how people – or AI algorithms – make judgments about their social environment (see Fig. 1). Whether such judgements occur in the context of assessing the quality of a job candidate, predicting the success of a stock on the stock market, or estimating another person's romantic interest, humans as well as AI algorithms often cannot observe the "true" state of the world directly. Instead, they rely on available information, or "cues", to infer a higher-level construct of interest. For example, they might consult a person's CV to scan for indicators of resilience, study prior stock performance in the context of current market trends or look for an overlap in interests that might indicate romantic compatibility. These observable cues function as a lens that mediates the human or AI's perception of the unobservable criterion.

Figure 1

Brunswik lens model applied to the perception of romantic attraction (criterion) by human and artificial agents (judgment) who can access observable indicators (social cues).



Importantly, the Brunswik lens model introduces several metrics that can be used to probe and evaluate LLMs' capacity to accurately decode social cues. First, cue validity represents the relationship between observable (social) cues and the unknown criterion, which is the target of the judgment. In the context of romantic attraction, the observable cue of physical attractiveness, for example, is known to be related to romantic attraction[21–24]. Second, cue utilization indicates the extent to which a human or AI algorithm uses a particular cue in their judgment of the target. For example, different agents might consider physical attractiveness more or less important in their prediction of romantic attraction, regardless of whether the cue itself is, in

fact, valid. Both cue validity – i.e. the presence of valid, observable cues – and cue utilization – i.e., the use of these cues by human or artificial judges – are necessary for accurate social judgments, which are captured by the accuracy metric of the lens model. Specifically, accuracy captures the correlation between the criterion (i.e., romantic attraction) and the prediction made by the human or artificial judge (i.e., predicted romantic attraction).

Our investigation relies on the findings of prior research suggesting that there are at least some valid, observable social cues of romantic attraction[21,22,25,26]. Notably, most this existing work has focused on physical attractiveness as a driver of dating choices[21,23,24] or investigated non-verbal (e.g. laughing, gaze[27–29]) or paraverbal behaviors (e.g. pitch[6,30]) that are unobservable to the purely text-based LLMs deployed in this paper. However, there is at least some evidence for verbal indicators of romantic attraction in the context of online dating[31] and existing romantic relationships[32–34]. Using the structural analysis of social behavior coding scheme (SASB[35]), for example, Eastwick and colleagues[2] identified several verbal indicators of smooth versus awkward dates. Participants on dates that were rated as smooth tended to speak more warmly and were more other-focused than those on awkward dates[2]. Similarly, romantic attraction has been linked to linguistic style matching (i.e., the extent to which people are similar in how they talk with one another[36]), as well as the frequency by which speed-dating participants asked follow-up questions[5]. Taken together, existing research suggests that there are at least some valid verbal indicators of romantic

attraction (i.e., some degree of cue validity) LLMs could draw on when judging the social dynamics of speed dates.

However, it remains an open question whether LLMs indeed utilize these – or other currently unknown — cues to romantic attraction in their social judgments. On the one hand, the ability of LLMs to process and analyze vast amounts of language data reflecting the human experience across a wide range of domains (e.g., social media posts, news articles, stories, and cultural explorations) makes them ideal candidates for the detection of social patterns. If LLMs are indeed able to detect cues of romantic attraction, this would not only support their potential for relational agents but also offer a promising path for better understanding romantic attraction itself. For example, it is possible that LLMs identify cues that are valid indicators but have been overlooked by previous investigations, which have more narrowly focused on theoretically meaningful, pre-defined cues.

On the other hand, most current LLMs still lack the ability to process crucial non-verbal and paraverbal cues—such as facial expressions, body language, tone of voice, and timing—that often play an important role in romantic interactions. Additionally, there is a risk that AI models merely perpetuate stereotypical patterns of romantic attraction that form the foundation of their training and overlook more idiosyncratic expressions of romantic preferences and expressions.

In this paper, we aim to empirically test LLMs' ability to detect social cues of romantic attraction Specifically, our analyses aim to answer three interrelated questions. Can LLMs like ChatGPT predict romantic interest from brief social interactions between humans? How does their predictive accuracy compare to that of human observers? Which linguistic and conversational cues do LLMs (and human judges) rely on in their predictions, and which among these cues are valid?

Methods

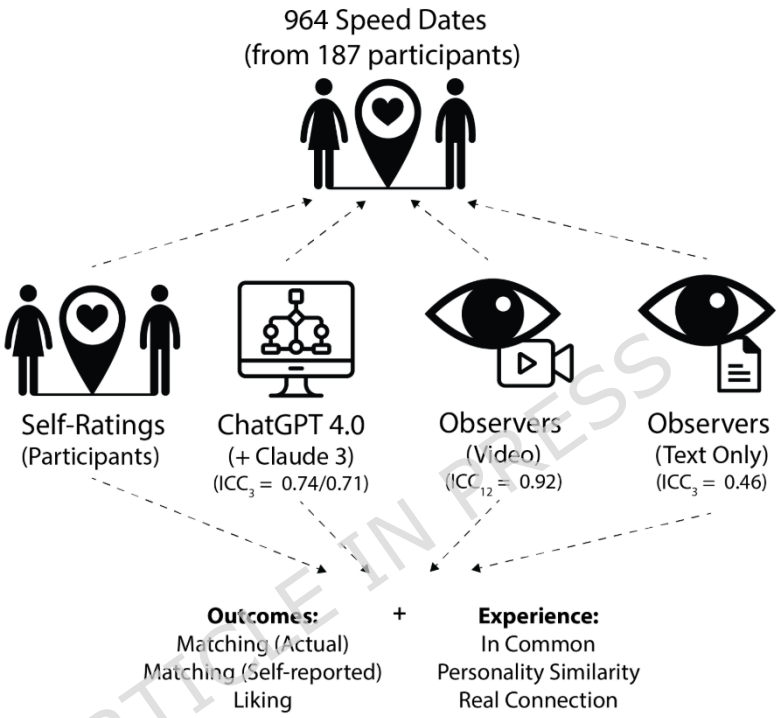
Participants and Procedure

The original dataset contained 1,039 speed dates from 187 undergraduate students (93 women and 94 men; age = 19.6±1.2 years old) who attended one of eight speed dating events at a large Midwestern University in 2007. Participants were recruited via flyers and e-mails and engaged in approximately 12 speed dates, each lasting four minutes and consisting of a man and a woman (see **Fig. 2** for an overview of the Study Design and Supplemental Material A in the SI for recruitment materials). Immediately following each speed date, participants evaluated the experience (see the Methods section for more details and refer to Finkel and colleagues[37] for the original publication of the dataset). The data collection for this study was approved by the Ethics Committee of Northwestern University (Protocol #1343-019). All research was performed in accordance with the University's

guidelines for human subjects research, and all participants provided informed consent before participating in the study.

Figure 2

Overview of study design.



Note. Speed dates were independently evaluated by four judges: speed date participants, ChatGPT (and Claude 3 in supplementary analyses supporting generalizability), human observers with access to video recordings of the speed dates, and observers with access to speed date transcripts only. All judges provided information on both the (predicted) speed-dating outcomes and experiences. ICC = Intraclass Coefficient (measure of inter-rater reliability).

All speed dates were recorded on video and transcribed by human professionals. We were able to match 964 of the transcripts based on their unique identifiers in the self-report data and file names of the transcripts, which included an average of 864±147 words per date. The sample size allows us to detect small effects of r = 0.10 with a power of 90% (assuming

an alpha of 0.05). The data and code needed to replicate our findings are openly available on the project's OSF page (view-only for peer review purposes).

Measures

The speed dates were evaluated by four independent sources: (i) the speed daters immediately after the interaction; (ii) ChatGPT (gpt-4-0613), which analyzed the written transcripts of the speed dates (the Supplementary Information also reports the findings for judgements made by Anthropic's Claude3, supporting the generalizability of our findings to other LLMs), (iii) eighteen judges (of which 12 raters scored every item), who watched the speed-dating videos; and (iv) four judges who read a random subset of 500 speed-dating transcripts without access to the videos.

All four sources rated the same set of five statements using a 9-point Likert scale (see Supplemental Materials B, C, and D in the SI for the specific items, the distributions of judgments, and ICCs). Two of the statements captured speed-dating *outcomes*: (i) participants' level of liking for their partner (e.g., ", and (ii) the likelihood of saying "yes" to a potential follow-up date (e.g., "I am likely to say yes to my interaction partner"). The other three captured speed daters' *experiences* as potential mechanisms: participants' beliefs that they (i) had "a lot in common" with their partner, (ii) had "similar personalities," and (iii) experienced "a real connection." ChatGPT and the human judges reading the transcripts rated the outcomes at the dyad level

(e.g. “On a scale from 1-9, how likely do you think the interaction partners will say “yes” to each other and exchange contact information?”). In contrast, both participants and the human judges with access to the video recordings rated them at the level of individual participants; To score the speed date as a whole, we averaged those individual ratings across both interaction partners.

While interrater reliabilities for the human judges with access to the speed-dating videos were high (average ICC = 0.92 across the two speed-dating outcomes and experiences), they were significantly lower for the human judges who only read the transcripts (average ICC = 0.46). Notably, the interrater reliabilities associated with the ratings produced by ChatGPT (and Claude 3 in the SI) can be influenced by adjusting the model’s temperature, a parameter that determines how similar the model’s responses are across different rounds[38]. For the purpose of this study, we kept the models’ standard settings (temperature=1) to test their baseline performance. The average agreement across the three independent API queries for ChatGPT was ICC=0.74 (average ICC = 0.71 for Claude 3). On average, the ratings of ChatGPT and Claude 3 were correlated by 0.71.

After the speed-dating event concluded, participants indicated whether they wanted to exchange contact information with each partner for a potential follow-up date. Among all the speed-dating dyads, 23% matched by exchanging contact information. This mutual signal of interest or “matching”

served as a third, objective speed-dating outcome with clear real-world consequences. Because the decision to match happened without the partner's immediate knowledge and was made several hours after the event, at home in private, actual matching should be less susceptible to potential biases than the previously outlined self-report measures (e.g. participants' desire to be polite, social pressure). Consequently, we use actual matching as our main metric of romantic attractions. Notably, however, the fact that our findings are robust across different subjective and objective indicators, increases our confidence in their validity as well as generalizability.

Analyses

We conducted two complementary analyses to evaluate ChatGPT's prediction accuracy. First, we assessed dyad-level predictive validity by examining the zero-order correlations between LLM-generated predictions, actual matching as well as the subjective speed-dating outcomes and experiences. This first set of analyses evaluates whether LLMs can infer attraction from conversational content at the level of individual interactions (i.e. answering the question "*Are ChatGPT's predictions aligned with real outcomes?")*, and as such is most closely aligned with the accuracy metric in the Brunswik Lens Model. However, because participants engaged in multiple dates, we also ran a series of mixed-effects regression models that account for the non-

independence of observations by modeling crossed random intercepts for both partners[39].

Results

Research Question 1: Can ChatGPT predict speed-dating outcomes and experiences?

To assess whether ChatGPT can detect linguistic signals of romantic attraction, we first evaluated predictive accuracy at the level of individual dates (i.e., accuracy in the Brunswik lens model). Specifically, we examined the association between ChatGPT's predictions and the objective outcome of mutual contact exchange. This analysis treats each dyad as an independent prediction target and provides a direct test of the model's ability to infer attraction from conversational content. To offer a comparison point, we evaluate ChatGPT's accuracy vis-à-vis that obtained by human judges, which include both the self-reported intentions of the speed-dating participants themselves as well as the post-hoc ratings made by human judges based on the speed-dating transcripts or videos (**Fig. 3A)**. The zero-order correlations of all variables and the equivalent results for Anthropic's Claude 3 (claude-3-opus-20240229) can be found in Supplemental Materials E and F in the SI.

Although ChatGPT's predictive accuracy of actual matching was relatively low (r=0.12, p<.001), the LLM performed on par with the human raters who had access to the same information (i.e., transcripts only; r=0.13, p=.003).

As expected, the human coders who could also draw on the non-verbal cues observable in the video recordings outperformed both ChatGPT and their human counterparts with less information (r=0.31, p< .001).

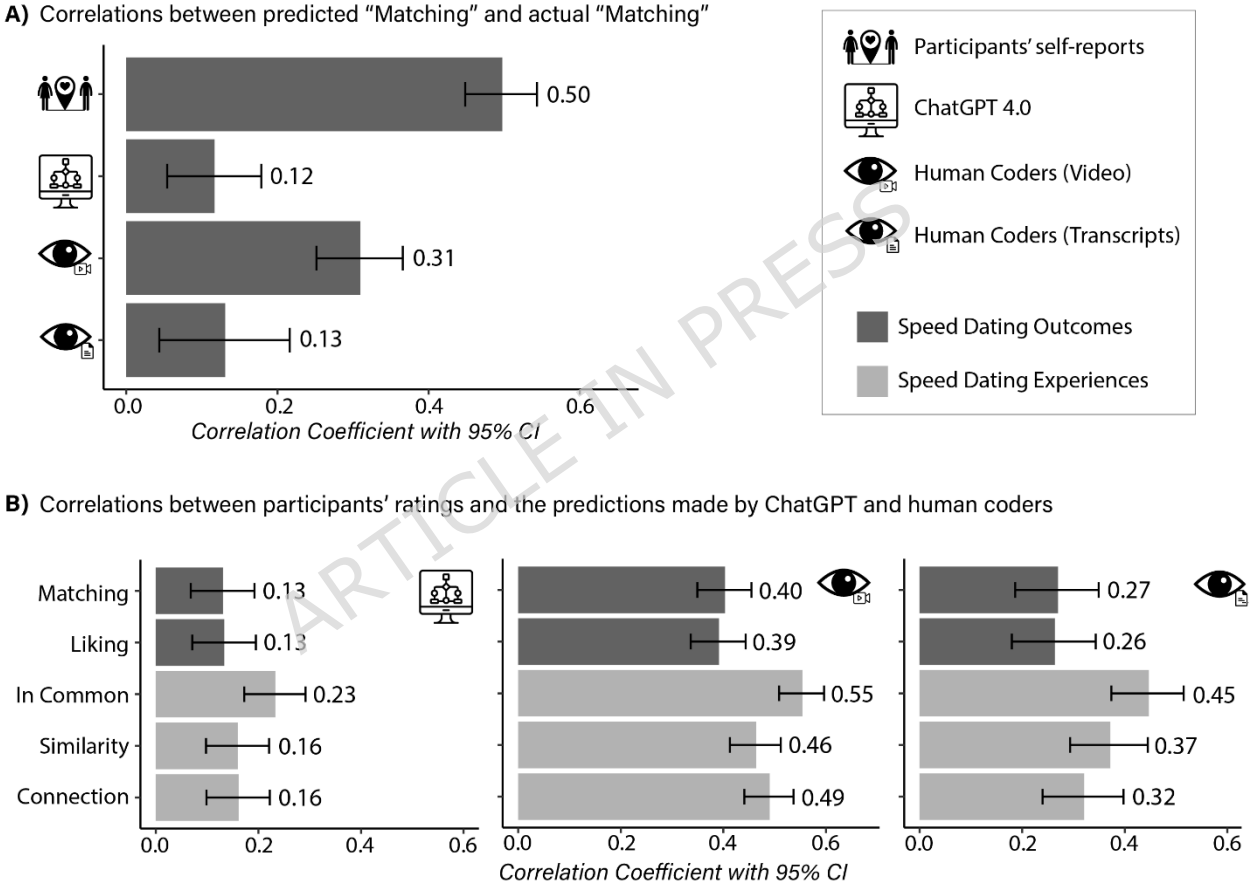
Notably, ChatGPT successfully predicted matching above and beyond participants' self-reported intentions (B=0.19, SE=0.09, z=2.17, p=.030). In contrast, we did not observe incremental predictive power for the ratings of the human judges who had access to transcripts only (B=0.03, SE=0.16, z=0.19, p=0.853). This suggests that ChatGPT observed unique social cues in the speed-dating transcripts that were neither accessible to the speed-dating participants themselves nor the coders who evaluated the same transcripts (note that human coders in the video condition did provide incremental predictive power: B=0.46, SE=0.10, z=4.62, p<.001.).

Next, we examined the ability of ChatGPT and human judges to predict participants' judgments of romantic attraction at the end of the speed date. ChatGPT predicted participants' ratings with correlations ranging between r = 0.13 to r = 0.23 (all p<.001; **Fig. 3B**), even after accounting for word count. The correlations were markedly lower than those made by human judges in both conditions (average correlation for observers with access to videos r = 0.46, and transcript only r = 0.33), suggesting that human judges reading the transcripts are better at predicting participants' self-reported experience of the speed date but not their actual commitment to exchanging contact information (actual matching). Notably, the relative levels of accuracy across

predicted metrics appeared largely consistent across all judges, with participants' more proximal experiences being easier to predict (in particular, whether they had a lot in common) than the more distal, overall evaluations of the speed dates.

Figure 3

Correlations between the speed-dating ratings of different evaluators.



Note. A) Point-biserial correlations between the behavioral speed-dating outcome of matching and the predictions made by different judges of whether participants would exchange contact information. B) Pearson correlations between participants' self-reported outcomes and experiences at the end of the speed date and the predictions made by ChatGPT and the human judges.

Because each participant engaged in multiple speed dates, observations were not statistically independent. To quantify the impact of individual differences on mutual matching, we first estimated a null crossed random-intercept logistic model with random intercepts for both daters. We then computed variance-partition coefficients on the latent (logit) scale to estimate the share of outcome variance attributable to male- and female-specific baseline matching propensity versus residual dyad-level variation (with residual variance fixed at $\pi^2/3$). The null model indicated that stable individual differences accounted for a roughly a third of the variance in mutual matching (34%), with most variability driven by residual, dyad-specific variance (see Table S3 in Supplementary Materials G for the results of variance decomposition across all subjective outcomes, which ranged between 12 and 21% of variance accounted for by individual differences). Even though the dyadic-specific variance includes both relationship-specific effects and noise, its large share suggests that matching outcomes are substantially influenced by interaction-dependent factors. In the next step, we estimated mixed-effect models for each outcome, including a fixed effect for ChatGPT's predictions and random intercepts for both interaction partners to account for the non-independence of observations. Once again, ChatGPT's predictions were associated with a greater likelihood of mutual matching in the logistic model (OR = 1.27, $p < .001$). Likewise, the linear mixed-effects models predicting the subjective outcomes showed that ChatGPT's predictions were also positively related to participants' self-

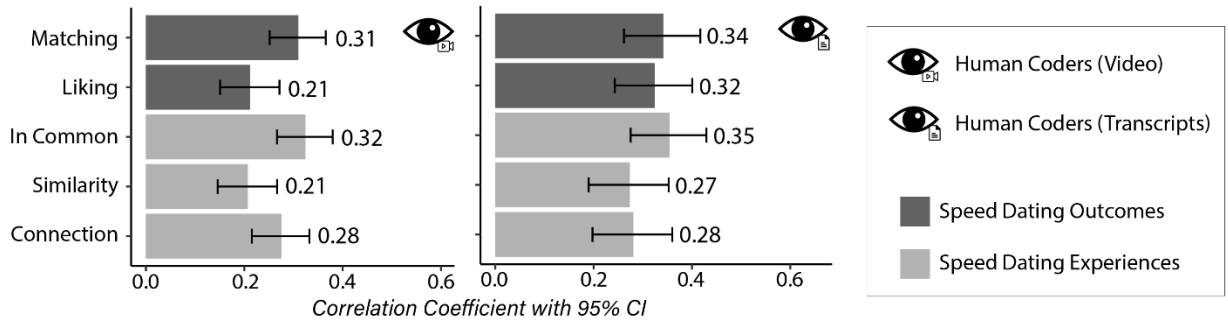
reported matching intentions (B = 0.13, $p = .005$), liking (B = 0.13, $p = .012$), perceived similarity (B = 0.24, $p < .001$), commonality (B = 0.30, $p < .001$), and connection (B = 0.16, $p < .001$). As for the raw correlations, the overall variance explained by the fixed effect of ChatGPT's prediction was modest in magnitude (marginal R^2s ≈ .01–.04; see Table S4 in Supplementary Material H for the full model output).

Research Question 2: Do ChatGPT's predictions overlap with those made by human judges?

In addition to comparing the predictive accuracies of ChatGPT and human raters, we also investigated the extent to which ChatGPT and human judges relied on similar cues when making their predictions. As the correlations in **Fig. 4** show, there was meaningful overlap between the predictions of ChatGPT and the human judges (mean $r = 0.29$ across all metrics and conditions). The fact that these correlations were higher than those observed with participants' self-reported scores (**Fig. 3B**) suggests that both ChatGPT and human judges utilize cues that are commonly considered markers of successful interactions but are not empirically associated with actual speed-dating outcomes or experiences (see the following results section for more details).

Figure 4.

Correlations between ChatGPT's predictions and the ratings made by human judges.



Although the predictions of ChatGPT and human judges overlap, they were not redundant in predicting speed-dating outcomes and experiences. For at least some of the metrics, both ChatGPT and human judges offered significant predictive power when added simultaneously into a linear regression model (notably, the effects of ChatGPT became non-significant when added with both, text-based and video-based, human judgments at the same time).

Research Question 3: Which linguistic and conversational cues are driving the predictions of ChatGPT, and are these cues valid indicators of romantic attraction?

To better understand how ChatGPT arrived at its predictions, we conducted two sets of additional analysis: First, we mapped ChatGPT's predictions against established markers of linguistic styles. Specifically, we used LIWC[40] – a widely used, dictionary-based text analysis tool – to categorize and quantify the psychological, emotional, and cognitive aspects of the speed-dating transcripts along 117 validated dictionaries. Second, we prompted ChatGPT to explain each of its predictions and subsequently asked it to synthesize these explanations in a parsimonious taxonomy of romantic

attraction markers. In both cases, we applied a Brunswik lens approach to test whether the cues utilized by the model (i.e., the language dimensions correlated with ChatGPT's predictions of attraction) were actually valid indicators of romantic attraction (i.e., exchange of contact information by participants).

LIWC Mapping

To explore the linguistic features associated with ChatGPT's predictions as well as actual matching, we extracted all 117 available LIWC dimensions for each of the transcripts. The dimensions cover a wide range of categories, such as linguistic dimensions (e.g., pronouns, verb tense), psychological constructs (e.g., affective processes, social concerns, cognitive processes), and personal concerns (e.g., work, money, health). We subsequently correlated the LIWC scores for each transcript with the predicted matching scores as well as actual matching outcome (see OSF for all correlations).

Table 1 displays the 10 LIWC dimensions (in addition to overall word count) that showed the highest absolute correlations with GPT's predicted likelihood of saying "yes.", capturing ChatGPT's cue utilization independent of accuracy. To increase the robustness of the relationships displayed we only included dimensions that showed consistent effects (in both direction and significance) in two randomly divided subsamples of the overall sample (see Supplemental Material I for the correlations in the subsamples). The examples suggest that GPT indeed relied on the affective valence of

conversations but also highlight the importance of other dimensions unrelated to emotional valence (e.g., high levels of certitude and a reduced focus on the present). In addition, the table captures cue utilization of human judges for the same dimensions as well as cue validity (i.e., correlation with the actual outcome). As the comparison between these different variables suggests there is a strong overlap between the cues utilized by ChatGPT and Humans, but not always high levels of accuracy (i.e., utilized cues are not always valid indicators). In some instances, the utilized cues are not only non-predictive of actual matching but, in fact, point in the opposite direction. For example, both ChatGPT and humans consider negations and the use of auxiliary verbs a negative indicator of romantic attraction, while both are indeed positively associated with actual matching.

Table 1.

LIWC dimensions that were most strongly correlated with ChatGPT's predictions of romantic attraction (i.e., cue utilization) alongside the cue utilization of human judges and cue validity.

Dictionary	Examples	Cue Utilization (ChatGPT)	Cue Utilization (Human)	Cue Validity
Negations	Not, no, never, nothing	-0.29	-0.06	0.09
Positive Tone	Good, well, new, love	0.28	0.19	0.05
Emotional Tone	Degree of positive (negative) tone	0.26	0.16	0.03
Affect	Good, well, happy, hope	0.26	0.20	0.06
Positive Emotion	Good, love, happy, hope	0.20	0.31	0.09

Word Count	Total word count	0.20	0.36	0.13
Emotion	Degree of positive (negative) emotion	0.19	0.28	0.08
Auxiliary Verbs	Is, was, be, have	-0.18	-0.12	0.06
Differentiation	But, not, if, or	-0.17	-0.11	-0.01
Present Focus	Is, are, I'm, can	-0.15	-0.14	0.02
Certitude	Really, actually, of course, real	0.14	0.08	-0.003

Note. Cue Utilization = correlation between judgement and cue, Cue Validity = Correlation between actual matching and cue.

To estimate the overall degree to which ChatGPT relied linguistic cues, we conducted a cross-validated out-of-sample estimation of the variance in ChatGPT's judgments of mutual attraction (i.e., the exchange of contact information) that could be explained by a linear combination of the 117 LIWC features. To do so, we first randomly selected 50% of the data (training dataset with 499 speed dates) to train a LASSO model, tuning the model parameter lambda in a 10-fold cross-validation. We subsequently applied the model to the remaining 50% of the data (testing dataset with 461 speed dates). The out-of-sample validation of the model on the testing data suggests that LIWC accounted for only ~12% of the variance in ChatGPT's predictions (r=0.35). This highlights that only a relatively small proportion of the variance in the predictions made by ChatGPT can be explained by the relatively comprehensive set of content dimensions included in the LIWC

dictionaries. The following sections therefore explore whether and to what extent ChatGPT might have relied on more complex conversational dynamics when judging romantic attraction.

ChatGPT-based Explanations

To probe the conversational cues underlying ChatGPT's predictions, we move beyond pre-defined linguistic markers and, instead, introduce a novel approach that leverages ChatGPT's ability to reflect on its own reasoning. That is, we prompted the model to explain the predictions it made through a zero-shot chain-of-thought process[41]. Specifically, we prompted ChatGPT: "How likely do you think the interaction partners will say "yes" to each other and exchange contact information? Let's think step by step." (see our OSF page for all generated explanations). A manual inspection of the explanations revealed that ChatGPT referred to many of the same indicators of romantic attraction previously identified in the literature (e.g. smoothness of conversation[2,42]). Instead of imposing a pre-existing taxonomy on the explanations generated by ChatGPT, we asked the model to "*Create a parsimonious taxonomy of explanation categories that captures recurring explanations and are psychologically meaningful*" (see Supplemental Material J in the SI for full details). Given the limited number of tokens per API query, we limited ChatGPT to the first 100 explanations.

The proposed taxonomy included eight broad categories (e.g., Conversation Flow and Engagement), and 19 secondary indicators (e.g., Smoothness of

Conversation, Mutual Engagement, Humor and Playfulness). Before proceeding with additional analyses, we excluded three categories we reasoned would be difficult to rate reliably based on textual data alone (i.e., External Factors and Contextual Influences, Internal States and Perceptions, Non-Verbal and Unspoken Cues). Notably, the inclusion of these categories in the proposed taxonomy could be explained in two ways: First, it is possible that ChatGPT complemented its analysis of the specific with its generic representation of romantic attraction. Second, many explanations mentioned the categories and its indicators as important drivers that could *not* be considered given the purely textual input (e.g. body language). Although these explanations explicitly highlight the absence of these cues from the transcripts, it is possible that ChatGPT nonetheless included them in the creation of its overall taxonomy. The final taxonomy with primary categories and secondary indicators is displayed in Table 3 (see Supplemental Material K in the SI for additional explanations of the indicators and excluded categories). Importantly, the proposed taxonomy is a bottom-up, exploratory framework that serves the particular purpose of exploring ChatGPT's decision-making. As such, it represents a reconstruction of the cues the model appears to rely on when making predictions (i.e., cue utilization), rather than an independently validated set of empirically established attraction markers (we examine the validity of these indicators in subsequent analyses). In other words, the taxonomy not meant as a definitive, exhaustive taxonomy

capturing all possible or plausible cues of romantic attractions or a canonical structure for how any such cues should be organized.

Table 3

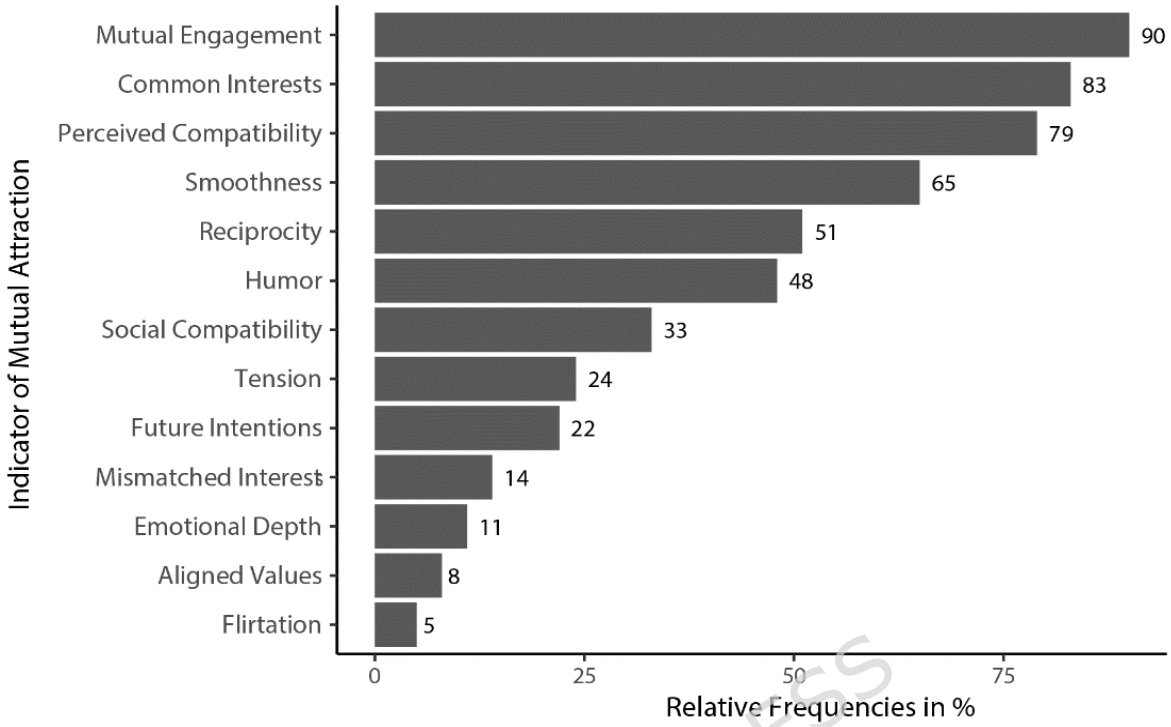
Bottom-up taxonomy of indicators of romantic attraction generated by ChatGPT

Primary Category	Secondary Indicator
Conversation Flow and Engagement	Smoothness of Conversation
	Mutual Engagement
	Humor and Playfulness
Shared Interests and Values	Common Interests
	Aligned Values and Goals
Emotional and Personal Connection	Emotional Depth
	Reciprocity of Disclosure
	Flirtation and Romantic Signals
Perceived Compatibility and Future Potential	Perceived Compatibility
	Expression of Future Intentions
	Social Compatibility
Conversational Challenges	Moments of Tension or Disagreement
	Mismatch in Interests or Values

As a first step in exploring the content of ChatGPT's explanations, we calculated the relative frequencies by which each of the indicators was mentioned as part of the explanations (see Supplemental Material L in the SI for details on how we extracted the frequencies). As Figure 5 shows, mutual engagement, common interests, perceived compatibility and smoothness of conversation were among the most commonly referenced explanations, followed by reciprocity and humor.

Figure 5

Relative frequencies of indicators of romantic attraction mentioned in ChatGPT's explanations of its predictions across all 964 speed dates.



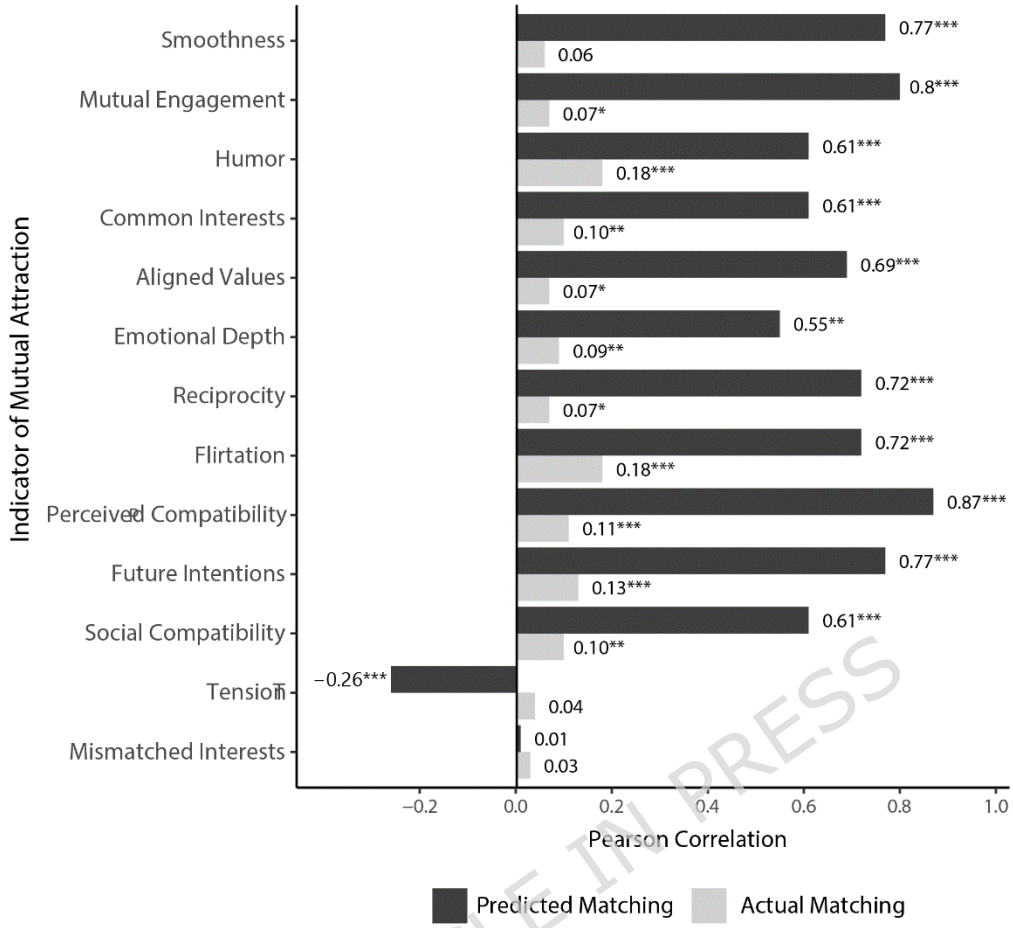
Next, we tested whether the explanations provided by ChatGPT were indeed related to its predictions of romantic attraction as well as actual matching. For this purpose, we prompted ChatGPT to rate each transcript on the thirteen indicators as well as the likelihood of matching (e.g. "*On a scale from 1-9, how smoothly did the conversation progress without awkward pauses or forced topics?*", see Supplemental Material M in the SI for the full prompt). As before, we repeated this exercise three times and calculated the average rating across all three rounds (average ICC = 0.61). Figure 6 displays the correlations between the thirteen indicators of romantic attracted and predicted matching scores (dark grey bars). Supporting the validity of ChatGPT's explanations, all except one indicator (i.e., Mismatched Interests) were significantly correlated with its predictions of matching, with all

indicators together explaining 80% of the variance in ChatGPT's predictions (following the same out-of-sample validation procedure as we described for the LIWC analysis)

Similar to the Brunswik lens approach applied to the LIWC dimensions, we further tested which of the utilized cues were indeed valid predictors of actual matching. To do so, we calculated the correlations between indicator ratings and actual matching (i.e., participants exchanging contact information). As Figure 6 shows, most indicators used by ChatGPT (i.e., 10 out of 13) were significantly associated with actual matching (light grey bars). These include Humor, Flirtation, Future Intentions, Perceived Compatibility, Common Interests, Social Compatibility, Emotional Depth, Aligned Values, Reciprocity and Mutual Engagement, in order of importance. Notably, when submitted to a linear logistic regression predicting actual matching, all ChatGPT-rated indicators combined explain approximately twice as much variance (based on the adjusted R-squared) as the direct matching prediction made by ChatGPT. This highlights the potential value of integrating theoretical depth into machine-based predictions.

Figure 6

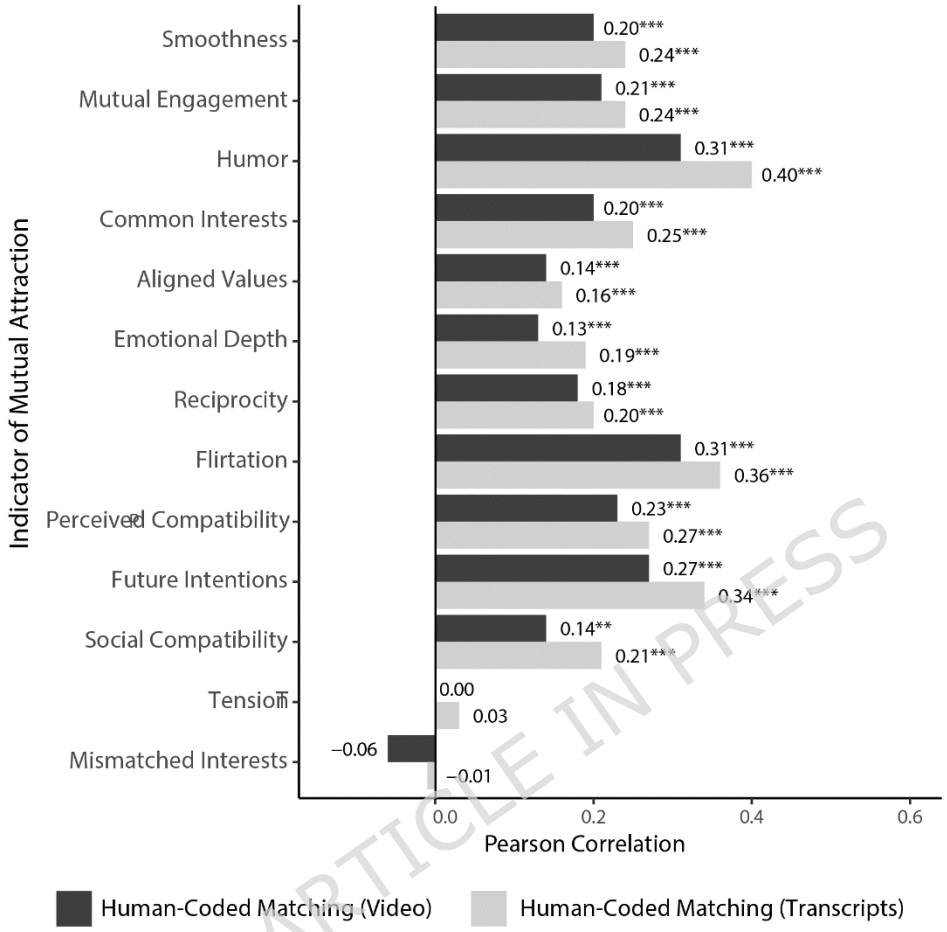
Correlations between indicators of romantic attraction as rated by ChatGPT as well as predicted matching (dark grey bars) and actual matching (light grey bars).



Finally, we tested the indicators as generated and rated by ChatGPT against the benchmark of human judgements. That is, we correlated the indicator ratings with the matching judgements made by the human observers with access to the speed-dating videos as well as transcripts only. Mirroring the previous findings in overlap between ChatGPT-based predictions and human ratings (see Figure 4), the correlations displayed in Figure 7 suggest that the human judges implicitly used many of the same cues used by ChatGPT.

Figure 7

Correlations between indicators of romantic attraction as rated by ChatGPT as well as human-coded matching of observers with access to the full video (dark grey) and transcripts only (light grey).



Discussion

Taken together, our findings suggest that LLMs can detect at least some of the conversational dynamics that lead strangers to "click" during romantic getting-to-know-you interactions. Notably, the accuracies were modest in magnitude (r=0.12-0.23). Yet, when predicting whether two people would eventually exchange contact information, the predictions made by ChatGPT were not only on par with those of human judges who had access to the same

information (transcripts) but incremental to participants' own ratings at the end of the speed date. Moreover, adding additional theory-driven inferential steps that focus the LLM's attention on observable, conversational indicators of romantic attraction led to a boost in the model's accuracy, highlighting the potential value of informing predictive approaches with theoretical frameworks and insights.

ChatGPT's predictions were also correlated with the predictions made by human raters (r=0.21-0.35), with both humans and ChatGPT relying on similar linguistic and conversational cues when making their judgements. Notably, ChatGPT and human raters were aligned in their utilization of both cues that were, indeed, valid indicators of romantic attraction and those that were not (see Brunswik's lens model[20]). This pattern suggests a reliance on shared but imperfect heuristics when judging romantic attraction, underscoring the potential risk of LLMs perpetuating stereotypical patterns of romantic attraction (observed in their training data) while overlooking more idiosyncratic and novel indicators of romantic preferences and expressions.

Our analysis exploring the linguistic and conversational features utilized by ChatGPT suggest that most of the variance in the model's predictions is explained by more complex conversational dynamics (e.g. the degree of mutual engagement or use of humor) rather than concrete linguistic markers (80% versus 12% of variance explained in an out-of-sample validation).

Moreover, our investigation of the specific linguistic and conversational features utilized by ChatGPT helps explain why the predictive accuracies remain relatively low. As our comparison between utilized and valid cues (see Tables 1 as well as Figure 6) suggests, there are several instances where ChatGPT – similar to human observers –used cues that are invalid. A better understanding of these mistakes could eventually help improve predictive accuracy in a modeling approach that leverages both theory development as well as data-driven pattern recognition.

In addition, our findings highlight the challenging nature of predicting romantic attraction. As our results suggest, ChatGPT – as well as human raters – were generally more accurate at predicting subjective attraction ratings than actual matching outcomes. This finding aligns with previous research showing that momentary feelings of attraction do not always translate into concrete dating decisions[25], underlining the complexity of romantic attraction judgments (as well as their underlying decision-making processes on the actor side) as a multistep and multifaceted process. Instead of being a monolithic concept, romantic attraction involves both immediate subjective experiences and additional context-driven considerations which highlights the importance of differentiating between the prediction of “in-the-moment” romantic attraction and the more challenging task of predicting the actual pursuit of a romantic partnership.

Contributions and Future Research

Our findings offer three core contributions. First, we extend the growing literature highlighting LLMs' ability to solve tasks that were previously considered to be the exclusive domain of human agents[9,15,43,44]. Prior work had shown how LLMs are capable of interpreting and mimicking the behavior of individuals, such as predicting people's personality from their social media profiles or short free-flow conversations[14,15], creating persuasive content that is customized to the psychological characteristics of their counterparts and often more persuasive than that generated by humans[43,45], or solving tasks that require theory of mind[9] as well as inductive reasoning[46]. Our work expands this literature by highlighting the ability of LLMs to capture social dynamics that play out between individuals in natural interactions.

Second, our findings offer novel insights into the verbal indicators of romantic attraction, supporting the proposition that verbal exchanges alone – without the observability of nonverbal or paraverbal cues – can reveal meaningful signals of romantic interest (i.e., both LLMs and human raters were able to predict actual matching at above-chance levels based on speed-dating transcripts alone). While most prior research has focused on static cues such as physical attractiveness[21,23,24], or nonverbal cues such as laughing or the gazes of dating partners[27–29], our research highlights the importance of more dynamic conversational features such as mutual engagement, shared interests or the reference to social activities. In doing so, our findings also go beyond existing work on verbal indicators of romantic attraction which has either focused on very specific mechanisms of language-

based interaction (e.g. language style matching[36]) or the overall quality of the interactions[47].

Finally, our research contributes to a better understanding of how LLMs (and humans) integrate available social cues into judgments about romantic attraction. Following a Brunswik lens approach, we showed how the predictions made by ChatGPT and human judges rely on common cues that are not necessarily valid predictors of romantic attraction. That is, while some of the cues utilized were indeed related to participants exchanging contact information after the speed date (e.g. references to emotions and social behaviors as well as perceived humor, flirtation or social compatibility; see the first column of Table 1 as well as Fig. 6), many of the cues utilized were not empirically related to matching (e.g., markers of certitude, future focus and insight as well as perceived smoothness or tension; see the second column of Table 1 as well as Fig. 6). Our findings align with existing research, demonstrating how people's lay theories of what inspires romantic attraction often differ from what actually inspires attraction[42].

Limitations

It has not escaped our notice that our study has several important limitations that should be addressed by future research. First, as speed dates are (by definition) brief, it is possible that LLMs might yield more accurate judgments when observing longer initial interactions. This is particularly true as speed dates might be rather similar in structure and content and remain more

superficial than other types of conversations due to their introductory nature (i.e., most people will start with brief introductions and small talk). Future research should investigate different types of social interactions (e.g. couples' text messages) and explore whether the accuracy of predictions can be improved by analyzing longer text excerpts.

Second, our analyses focused exclusively on short-term success metrics and proxies of romantic interest, such as self-reported liking or the exchange of contact information, which are likely influenced by factors such as politeness, uncertainty, reciprocity expectations, or perceived social desirability. While the combination of objective behavioral indicators with subjective self-reports increases the robustness of our findings, future research should investigate the ability of LLMs and human judges to predict medium to longer-term indicators of romantic interest (e.g., the likelihood of entering into a relationship and relationship satisfaction) and explore how attraction judgements change during the early stages of a developing relationship. A particularly interesting period to study might be the transition from initial interest to sustained romantic commitment[48], which requires romantic partners to incorporate new information and reevaluate prior impressions in response to changing relational contexts[49]. Akin to the romantic partners themselves, external judges (including LLMs) might need to update their use of different social cues across different stages. For instance, while an initial attraction judgment might be influenced by observable surface traits, later decisions – such as whether to pursue exclusivity or deepen emotional

intimacy – may depend on different information or a more nuanced integration of behavioral patterns and compatibility cues[25,49,50]. From a lens-model perspective, this means that individuals not only rely on new information but also reinterpret and weigh existing cues as the relationship progresses. What seemed attractive in a first encounter may take on a different weight when viewed in the context of everyday relationship experiences. Being able to keep up with the shifting meanings of social cues and the corresponding needs of their human counterparts, is likely to play a critical role in allowing relational agents to establish and maintain meaningful long-term relationships that develop from relatively superficial encounters to deeper bonds of trust.

Third, our analysis predominantly focused on predicting the average level of romantic attraction per dyad alongside the mutual decision to exchange contact information. However, as the Social Relations Model (see Social Relations Model[51,52]) highlights, romantic attraction is made up of three distinct elements: actor effects (i.e., how much a person generally desires others), partner effects (i.e., how desirable that person is in general), and relationship effects (whether any two individuals will be attracted to one another). While we explored some of these nuances by including the results of hierarchical models that account for individual differences (i.e. actor and partner effects) when estimating the relationship between ChatGPT's predictions and actual outcomes, future research should investigate these dynamics in more depth by treating each person as an individual datapoint.

Fourth, our sample consists of college students from a single University, most of which are in their early 20s and were born and/or raised in the U.S. Because LLMs were predominantly trained on the text generated by users in Western, English-speaking countries (the majority of which are White), the models might be better at predicting social dynamics for those populations than others. Hence, although ChatGPT was not necessarily trained for students (which might make the model worse for this particular population than others), it could still perform better in our sample than in other samples that are less represented in the corpora of text ChatGPT was trained on. In addition, our data was collected in 2007. Although textual data pre-2007 is heavily represented in the training data of LLMs like ChatGPT – e.g. in the form of books, news archives or encyclopedic writing – more informal, conversational data at internet scale is likely underrepresented compared to more recent time periods. While most drivers of romantic attraction are likely to be relatively stable across time (e.g. perceived similarity), it is possible that some romantic communication norms – e.g. norms related to gender, boundary-setting or humor – have changed over time. Notably, such a drift would make the predictive accuracies presented in this paper conservative estimates of ChatGPT's true capacity.

Fifth our study exclusively relied on language models and conversation transcripts, which omit crucial information such as body language and paralinguistic cues. The importance of such cues for predictions of romantic attractions is highlighted by the different levels of accuracy between human

raters with and without access to non-verbal cues. Specifically, human raters who not only read through the transcripts – analogous to the ChatGPT condition – but also watched the video recordings of the speed dates, showed significantly higher levels of both consistency in their ratings (i.e., higher levels of inter-rater agreement) as well as predictive accuracy. This discrepancy likely stems from the fact that text-based media are low in richness and lack nonverbal and paralinguistic cues. As a result, the judgements made by third-party observers – both humans and LLMs – based on speed-dating transcripts alone are subject to higher levels of ambiguity and subjectivity[53–55], which, in turn, limits both their reliability and validity. In contrast, the video recordings contain non- and paraverbal cues that can remove ambiguity and decrease the risk of misinterpretation (e.g. a sarcastic statement might be much easier to identify when hearing the participants' voices and seeing their facial expressions). In addition, the non-verbal cues observable in the full recordings likely hold additional predictive signals that are not captured in the transcripts but are highly predictive of romantic attraction (e.g. smiling, leaning-in toward the partner). As generative AI continues to expand in the domains of computer vision and multimodal interfaces (e.g., OpenAI's GPT-4o model, which processes both spoken language and images), future research should test whether artificial agents can reach – or even surpass – the level of accuracy obtained by the human raters with access to the speed-dating videos. This technological advancement will allow for a deeper investigation of the interplay between

verbal and nonverbal signals of romantic attraction. While our study highlights the importance of verbal behavior, it is well-established that social impressions are often shaped by an interplay between multiple behavioral channels[56,56,57]. As LLMs develop a capacity to interpret social cues at different levels, they could be used to explore how the integration of different channels improves predictive accuracy and investigate the impact of mismatched verbal and nonverbal signals (e.g., the impact of enthusiastic language paired with disengaged body language). Investigating such discrepancies may help explain why some interactions appear promising on a verbal level but, in reality, do not lead to genuine romantic interest.

Conclusion

Taken together, our work suggests that LLMs like ChatGPT have some capacity to detect social dynamics in natural conversations, but that their overall predictive power remains modest. However, we consider our findings a conservative estimate of AI's capacity to understand human interactions and predict that its ability to interpret conversational dynamics on a much more holistic level is poised to increase significantly over the coming years.

Declarations:

- No funding was received for this study.
- The authors declare no conflict of interest.
- The project has received ethical approval from the Ethics Committee of [Blinded for Review] (Protocol #1343-019)
- The manuscript was prepared without the assistance of AI tools.
- Data and analyses codes are available on OSF

References

1 Berscheid E, Hatfield E. *Interpersonal Attraction*. Addison-Wesley, 1969.

2 Eastwick PW, Saigal SD, Finkel EJ. Smooth Operating: A Structural Analysis of Social Behavior (SASB) Perspective on Initial Romantic Encounters. *Soc Psychol Personal Sci* 2010; 1: 344–352.

3 Place SS, Todd PM, Zhuang J, Penke L, Asendorpf JB. Judging romantic interest of others from thin slices is a cross-cultural ability. *Evol Hum Behav* 2012; 33: 547–550.

4 Gottman JM, Coan J, Carrere S, Swanson C. Predicting Marital Happiness and Stability from Newlywed Interactions. *J Marriage Fam* 1998; 60: 5–22.

5 Huang K, Yeomans M, Brooks AW, Minson J, Gino F. It doesn't hurt to ask: Question-asking increases liking. *J Pers Soc Psychol* 2017; 113: 430–452.

6 McFarland DA, Jurafsky D, Rawlings C. Making the Connection: Social Bonding in Courtship Situations. *Am J Sociol* 2013; 118: 1596–1649.

7 Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN *et al.* Attention is All you Need. In: *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2017https://papers.nips.cc/paper_files/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html (accessed 30 Mar2026).

8 Jakesch M, Hancock JT, Naaman M. Human heuristics for AI-generated language are flawed. *Proc Natl Acad Sci* 2023; 120: e2208839120.

9 Kosinski M. Evaluating large language models in theory of mind tasks. *Proc Natl Acad Sci* 2024; 121: e2405460121.

10 Lei S, Dong G, Wang X, Wang K, Qiao R, Wang S. InstructERC: Reforming Emotion Recognition in Conversation with Multi-task Retrieval-Augmented Large Language Models. 2024. doi:10.48550/arXiv.2309.11911.

11 Wang X, Li X, Yin Z, Wu Y, Liu J. Emotional intelligence of Large Language Models. *J Pac Rim Psychol* 2023; 17: 18344909231213958.

12 Zhang W, Deng Y, Liu B, Pan SJ, Bing L. Sentiment Analysis in the Era of Large Language Models: A Reality Check. 2023. doi:10.48550/arXiv.2305.15005.

13 Zhang Z, Peng L, Pang T, Han J, Zhao H, Schuller BW. Refashioning Emotion Recognition Modeling: The Advent of Generalized Large Models. *IEEE Trans Comput Soc Syst* 2024; 11: 6690–6704.

14 Peters H, Cerf M, Matz SC. Large Language Models Can Infer Personality from Free-Form User Interactions. 2024. doi:10.48550/arXiv.2405.13052.

15 Peters H, Matz SC. Large language models can infer psychological dispositions of social media users. *PNAS Nexus* 2024; 3: pgae231.

16 Ko D, Lee S, Kim G. Can Language Models Laugh at YouTube Short-form Videos? 2024. doi:10.48550/arXiv.2310.14159.

17 Liu H, Wang W, Li H. Towards Multi-Modal Sarcasm Detection via Hierarchical Congruity Modeling with Knowledge Enhancement. In: Goldberg Y, Kozareva Z, Zhang Y (eds). *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics: Abu Dhabi, United Arab Emirates, 2022, pp 4995–5006.

18 Neuman Y, Cohen Y. A Dataset of 10,000 Situations for Research in Computational Social Sciences Psychology and the Humanities. *Sci Data* 2023; 10: 505.

19 Yarkoni T, Westfall J. Choosing Prediction Over Explanation in Psychology: Lessons From Machine Learning. *Perspect Psychol Sci* 2017; 12: 1100–1122.

20 Brunswik E. Historical and Thematic Relations of Psychology to Other Sciences. *Sci Mon* 1956; 83: 151–161.

21 Back MD, Penke L, Schmukle SC, Sachse K, Borkenau P, Asendorpf JB. Why Mate Choices are not as Reciprocal as we Assume: The Role of Personality, Flirting and Physical Attractiveness. *Eur J Personal* 2011; 25: 120–132.

22 Eastwick PW, Finkel EJ, Joel S. Mate evaluation theory. *Psychol Rev* 2023; 130: 211–241.

23 Montoya RM. I’m Hot, So I’d Say You’re Not: The Influence of Objective Physical Attractiveness on Mate Selection. *Pers Soc Psychol Bull* 2008; 34: 1315–1331.

24 Walster E, Aronson V, Abrahams D, Rottman L. Importance of physical attractiveness in dating behavior. *J Pers Soc Psychol* 1966; 4: 508–516.

25 Asendorpf JB, Penke L, Back MD. From Dating to Mating and Relating: Predictors of Initial and Long–Term Outcomes of Speed–Dating in a Community Sample. *Eur J Personal* 2011; 25: 16–30.

26 Joel S, Eastwick PW, Finkel EJ. Is Romantic Desire Predictable? Machine Learning Applied to Initial Romantic Attraction. *Psychol Sci* 2017; 28: 1478–1489.

27 Grammer K, Honda M, Juette A, Schmitt A. Fuzziness of nonverbal courtship communication unblurred by motion energy detection. *J Pers Soc Psychol* 1999; 77: 487–508.

28 Grammer K, Kruck K, Juette A, Fink B. Non-verbal behavior as courtship signals: the role of control and choice in selecting partners. *Evol Hum Behav* 2000; 21: 371–390.

29 Renninger LA, Wade TJ, Grammer K. Getting that female glance: Patterns and consequences of male nonverbal behavior in courtship contexts. *Evol Hum Behav* 2004; 25: 416–431.

30 Pisanski K, Oleszkiewicz A, Plachetka J, Gmiterek M, Reby D. Voice pitch modulation in human mate choice. *Proc R Soc B Biol Sci* 2018; 285: 20181634.

31 Lee AJ, Jones BC, DeBruine LM. Investigating the association between mating-relevant self-concepts and mate preferences through a data-driven analysis of online personal descriptions. *Evol Hum Behav* 2019; 40: 325–335.

32 Bowen JD, Winczewski LA, Collins NL. Language Style Matching in Romantic Partners' Conflict and Support Interactions. *J Lang Soc Psychol* 2017; 36: 263–286.

33 Karan A, Rosenthal R, Robbins ML. Meta-analytic evidence that we-talk predicts relationship and personal functioning in romantic couples. *J Soc Pers Relatsh* 2019; 36: 2624–2651.

34 Simmons RA, Gordon PC, Chambless DL. Pronouns in Marital Interaction: What Do "You" and "I" Say About Marital Health? *Psychol Sci* 2005; 16: 932–936.

35 Dai M, Robbins R. Exploring the influences of profile perceptions and different pick-up lines on dating outcomes on tinder: An online experiment. *Comput Hum Behav* 2021; 117: 106667.

36 Ireland ME, Slatcher RB, Eastwick PW, Scissors LE, Finkel EJ, Pennebaker JW. Language Style Matching Predicts Relationship Initiation and Stability. *Psychol Sci* 2011; 22: 39–44.

37 Finkel EJ, Eastwick PW, Matthews J. Speed-dating as an invaluable tool for studying romantic attraction: A methodological primer. *Pers Relatsh* 2007; 14: 149–166.

38 OpenAI, Achiam J, Adler S, Agarwal S, Ahmad L, Akkaya I *et al.* GPT-4 Technical Report. 2024. doi:10.48550/arXiv.2303.08774.

39 Raudenbush SW, Bryk AS. *Hierarchical Linear Models: Applications and Data Analysis Methods*. SAGE, 2002.

40 Boyd RL. The Development and Psychometric Properties of LIWC-22. 2022.

41 Kojima T, Gu S (Shane), Reid M, Matsuo Y, Iwasawa Y. Large Language Models are Zero-Shot Reasoners. *Adv Neural Inf Process Syst* 2022; 35: 22199–22213.

42 Eastwick PW, Sparks J, Finkel EJ, Meza EM, Adamkovič M, Adu P *et al.* A worldwide test of the predictive validity of ideal partner preference matching. *J Pers Soc Psychol* 2024; 128: 123–146.

43 Bai H. Artificial Intelligence Can Persuade Humans on Political Issues. 2023.

44 Orrù G, Piarulli A, Conversano C, Gemignani A. Human-like problem-solving abilities in large language models using ChatGPT. *Front Artif Intell* 2023; 6. doi:10.3389/frai.2023.1199350.

45 Matz SC, Teeny JD, Vaid SS, Peters H, Harari GM, Cerf M. The potential of generative AI for personalized persuasion at scale. *Sci Rep* 2024; 14: 4692.

46 Bhatia S. Inductive reasoning in minds and machines. *Psychol Rev* 2024; 131: 1373–1391.

47 Sprecher S, Duck S. Sweet Talk: The Importance of Perceived Communication for Romantic and Friendship Attraction Experienced During a Get-Acquainted Date. *Pers Soc Psychol Bull* 1994; 20: 391–400.

48 Campbell L, Stanton SCE. The Predictive Validity of Ideal Partner Preferences in Relationship Formation: What We Know, What We Don’t Know, and Why It Matters. *Soc Personal Psychol Compass* 2014; 8: 485–494.

49 Eastwick PW, Finkel EJ, Eagly AH. When and why do ideal partner preferences affect the process of initiating and maintaining romantic relationships? *J Pers Soc Psychol* 2011; 101: 1012–1032.

50 Campbell L, Lackenbauer SD, Muise A. When Is Being Known or Adored by Romantic Partners Most Beneficial? Self-Perceptions, Relationship Length, and Responses to Partner's Verifying and Enhancing Appraisals. *Pers Soc Psychol Bull* 2006; 32: 1283–1294.

51 Back MD, Kenny DA. The Social Relations Model: How to Understand Dyadic Processes. *Soc Personal Psychol Compass* 2010; 4: 855–870.

52 Kenny DA, La Voie L. The Social Relations Model. In: *Advances in Experimental Social Psychology*. Academic Press, 1984, pp 141–182.

53 Kiesler S, Siegel J, McGuire TW. Social psychological aspects of computer-mediated communication. *Am Psychol* 1984; 39: 1123–1134.

54 Daft RL, Lengel RH. Organizational Information Requirements, Media Richness and Structural Design. *Manag Sci* 1986; 32: 554–571.

55 Walther JB, Burgoon JK. Relational Communication in Computer-Mediated Interaction. *Hum Commun Res* 1992; 19: 50–88.

56 Breil SM, Osterholz S, Nestler S, Back MD. Contributions of nonverbal cues to the accurate judgment of personality traits. In: *The Oxford Handbook of Accurate Personality Judgment*. Oxford University Press, 2021, pp 195–218.

57 Back MD, Nestler S. Dual-Process Approaches to Personality. In: *Reflective and Impulsive Determinants of Human Behavior*. Routledge, 2016.