

Social Chatbots and Character: Mind-Modelling, Mimetics, and Postdigital Misperception

BELL, Alice <<http://orcid.org/0000-0001-9737-4081>>

Available from Sheffield Hallam University Research Archive (SHURA) at:

<https://shura.shu.ac.uk/37931/>

This document is the Accepted Version [AM]

Citation:

BELL, Alice (2026). Social Chatbots and Character: Mind-Modelling, Mimetics, and Postdigital Misperception. Language and Literature. [Article]

Copyright and re-use policy

See <http://shura.shu.ac.uk/information.html>

Social Chatbots and Character: Mind-Modelling, Mimetics, and Postdigital Misperception

1 Introduction

Technological developments in and greater availability of generative AI has led to increased and wider use of chatbots across healthcare, commerce, retail and entertainment. Not only has this changed the way that humans access goods and services, but it has also reignited and recontextualised interest in the relationship between humans and machines across digital to non-digital boundaries, challenging perceptions of what is ‘real’ and what is not. A concern with the ontology of our contemporary experience is also reflected in the postdigital research paradigm. The ‘postdigital’ refers to a point in digitized societies in which ‘the historical distinction between the digital and non-digital becomes increasingly blurred’ (Berry and Deiter, 2014: 2) leading to a state in which the novelty of digital media has worn off and is simply commonplace (cf. Cramer, 2014). Thus, ‘we are increasingly no longer in a world where digital technology and media is separate, virtual, ‘other’ to a ‘natural’ human and social life’ (Jandrić et al., 2018: 893) but experiencing ‘an epistemological and pragmatic shift in everyday life towards the use of computational systems to support and mediate life itself’ (Berry and Deiter, 2014: 1). From a postdigital perspective, while materially distinct, our online experiences and related existence(s) are just as real as our offline experience and existence(s) with the distinction between them becoming increasingly less pronounced, affecting the way we think about, act, and communicate in and across the two domains (e.g. Lindgren, 2022).

In this article, I show how cognitive stylistics and narratology can contribute to our understanding and conceptualisation of user interaction with social chatbots – generative AI programs designed for social interaction with humans – as a manifestation of the postdigital. I report on the results of my empirical research on character.ai, one of the world’s largest social chatbot platforms. Using virtual ethnography, stylistics, and narratology I show how rhetorical and cognitive approaches to character can be operationalised to analyse user conceptualisation of chatbots but also how these approaches must be adapted for the networked, participatory, and generative context of AI. Focussing on responses to chatbot quality and ontology, I show how the human capacity for mind-modelling (Stockwell, 2009) is being used in a postdigital context, propose two new audience response types – transmimetic and hypermimetic – to account for the way that users evaluate social chatbots, and propose the new concept of ‘postdigital misperception’ to reflect the way that users of generative AI can misidentify the ontological origin of an LLM’s output.

2 AI, Social Chatbots, and Imitation

Known as bots, digital assistants and interactive agents, chatbots are a form of Artificial Intelligence that are now ubiquitously used in education, business, health, and entertainment (see Adamopoulou and Moussiades, 2020). Within the latter context, social chatbots are ‘artificial intelligence (AI) dialogue systems capable of having social and empathetic conversations with users’ (Brandtzaeg et al., 2022: 404). Thus, while all chatbots are designed to provide a conversational experience by responding to user input, social chatbots have a more specific interpersonal purpose. Rather than using predefined scripts, as might be found in some customer service bots, social chatbot platforms such as character.ai, Replika, and Snapchat’s

My AI use generative AI. Thus, they utilise a Large Language Models (LLMs) such that the chatbot generates linguistic content in response to a user's input by predicting which word is most likely to come next.

Existing research suggests users can engage in conversations with social chatbots that are fun, entertaining, educational, supportive, therapeutic, and even romantic or sexual. Skjuve et al.'s study of the social chatbot platform Replika finds that human to chatbot 'relationship development shares similarities with relationship development between humans' (2021: 2; cf. Schuetzler et al., 2014) and Brandtzaeg et al. that social chatbots 'mimic human emotions and language' (2022: 409) in order to encourage relationship development.

The capacity for imitating human communication speaks to some longstanding goals of AI in general. As Finlay and Dix (1996: 2) note, 'AI is concerned with building machines that can act and react appropriately, adapting their responses to the demands of the situation. Such machines should display behaviour comparable with that considered to require intelligence in humans'. AI research and development thus focusses on creating systems capable of performing tasks that typically require human intelligence such as learning, reasoning, problem-solving, recognizing patterns, and making decisions but, crucially, the machine output should be 'appropriate' to the context. Accordingly, different AI systems will be designed to meet the needs of different situations. As Caldarini et al. (2022: 1, my emphasis) explain in relation to chatbots specifically, they 'take natural language text as input, and the output should be the most *relevant* output to the user input sentence'. Like Finlay and Dix (1996), Caldarini et al. (2022) thus emphasise both the humanlike nature of chatbots but also the contextual suitability of the chatbot's response such that it makes the interaction feel appropriate, relevant, and authentic.

AI research has a long history of investigating ways in which computers might be able to mimic humans. Alan Turing's "imitation game" (1950) and Joseph Weizenbaum's (1966) interactive computer program "ELIZA" laid the foundations for investigating the extent to which human users might regard AI as human or humanlike. Natale (2021) shows how widespread and potentially invisible AI has since become. He proposes the term 'banal deception' to describe 'deceptive mechanisms and practices that are embedded in media technologies and contribute to their integration into everyday life [...] [with the use of banal used] to underline that these mechanisms are often taken for granted' (7). Natale ultimately suggests that humans are now so used to being deceived by AI in the form of, for example, voice assistants or customer service chatbots that they can forget that AI is only an imitation of human intelligence.

Existing research on chatbots shows that the capacity for imitating humans has been considered important for trust (e.g. Youn and Jin, 2021), user expectations (Go and Sundar, 2019) and audience retention (Zhang et al., 2024). Research has also shown that users of social chatbots can feel like chatbots are their friends and form parasocial romantic attachments with them (Brandtzaeg et al., 2022). Others have shown that chatbots can be unsuccessful in engaging users appropriately. As Adamopoulou and Moussiades (2020: 13) note 'chatbots often fail to recognize the intent of their interlocutor [...] [which] creates frustration for him/her' (cf. Caldarini et al., 2022). Skjuve et al.'s (2021: 9) empirical research on users of social chatbot platform Replika also finds that chatbots 'could fail to understand what the participants were saying or provide unintelligent or out-of-context answers'. Thus, despite significant developments in generative AI over recent years, there can be some incongruity between what the user expects and what the chatbot delivers, influencing a user's willingness to engage

and/or form a connection with this form of AI. Chen et al.'s (2024: 2) review of existing approaches to LLM-based role-playing suggests that 'character consistency, behavioral alignment, and overall attractiveness of the character portrayal' are important for 'immersive and believable character simulations that maintain continuity over interactions and adapt dynamically to dialogue contexts' (2). However, analyses of the LLMs' output are not provided.

Overall, there is some consensus that AIs including chatbots can imitate human communication and engage as well as retain users if they act appropriately. Users can also experience feelings towards an AI in the same way that they would a human being including various forms of kinship. What is missing from AI research is an understanding of the cognitive processes that users go through evaluating social chatbots. In what follows, I suggest that insights from narratological and stylistic approaches to character can provide this input.

3 Social Chatbots as Characters

Some existing research in narratology already shows how character theory can help our understanding of generative AI agents. Wilde (2025), for example, theorises users' perception of chatbot ontology, suggesting that they experience two conceptually distinct cognitive effects. First, taking influence from the findings of Weizenbaum's (1966) experiment, Wilde defines the Eliza-Effect as 'an emotional or affective engagement with technologies usually reserved for interactions with other human beings' (109). It can be used to explain the way that humans can form 'an intense *affective engagement* and an emotional bonding with the simple chatbot' (108). The second, which he derives from cognitive narratological approaches to character, is the Character-Effect: 'constructing or comprehending fictional (or, more broadly, represented) characters within a frame of make-belief' (110). This effect, Wilde shows, has been used within narratology more broadly to explain the way that audiences across media combine textual cues with their schematic knowledge about how humans think, behave, or feel to cognitively construct and relate to characters (cf. Culpepper, 2001; Schneider, 2001). While Wilde explains the Eliza- and Character-Effect as two distinct processes, he also suggests that engagement with LLM-based chatbots 'allows us not only to make believe a fictional (represented or implied) "continuous consciousness frame" (Palmer, 2010: 10) but also to search for traces of an actual one' (111) such that the two can become entangled. In relation to social chatbots specifically, Wilde suggests that 'we can see the "gap" between both effects [...] when generative platforms "roleplay" characters in LLM-based chatbot systems like Character.AI' (112). He speculates that social chatbot 'users distinguish between the LLMs as such and the roles (characters) they play within individual interactions' (112) such that they evaluate the humanlikeness of the chatbot and the ability of that anthropomorphized chatbot to play a particular character role separately.

Drawing on Frow's (2016) theory of character as 'quasi-person', Blom (2022) also argues for conceptualising AI agents as characters. She shows that Frow identifies a 'tension between thinking of characters as pieces of writing and imagining them as person-like entities' such that 'we tend to think of characters as person-like, but they have different modes of existing' (171). Blom's focus is voice assistants such as Siri and Alexa rather than social chatbots, but her approach shows how narratology can be used to understand engagement with AI more broadly. In particular, she suggests that because interaction between AI and user crosses an ontological boundary between real and virtual, 'users simultaneously regard voice assistants [...] as having human intentionality and tacitly understand that they are AI-driven entities' (173), a phenomenon she defines as 'double-consciousness'. She thus suggests that interactions that

take place across ontological boundaries mean that users can regard AI agents as potentially human and not just as non-human figures. Like Wilde, she thus sees the idea of both human-likeness and character construction as central for understanding audience's cognitive engagement with generative AI agents.

Bernini also shows how social chatbots can be considered characters, arguing for the benefits of using a narratological approach to *fictionality* thus focusing on the interplay of different ontological domains in the user-to-chatbot interaction. Like Wilde and Blom, he conceptualises user experience of chatbot interaction as a dual cognitive process, proposing the concept of 'fictive double awareness' to describe the sense that 'we simultaneously know and partially forget that we are the only human in the loop' (2025: 494-5). However, he also shows the importance of considering where users think chatbots exist in order to understand their overall ontological status. He suggests that users can be immersed by social chatbots and thus psychologically projected into a fictional world while also simultaneously experiencing 'emersivity': 'the sensed permeation of fictional elements into the real world of users' (468). Contrasting the interactive experience of chatbots with the reading of print fiction, he shows how a medium-specific approach is needed to understand the ontological affordances of generative AI agents and calls for empirical work in this area.

The preceding material shows the theoretical advancements made by applying narratological approaches to chatbots as characters. Crucially, unlike structuralist or so-called 'de-humanizing' approaches to character which 'insist that [characters] have a purely textual existence' (Culpepper, 2001: 7), Wilde, Blom, and Bernini draw on what have been called 'humanizing' (Culpepper, 2001) or 'humanistic' (Blom, 2022) approaches that consider the ways in which audiences draw on the same cognitive resources as they do when they encounter real people. Approaches that focus on the way in which audiences see characters as humanlike are particularly relevant for the analysis of social chatbots because, as the preceding section shows, they are designed to mimic human interaction and audiences tend to anthropomorphize them.

Outside of AI, cognitive and rhetorical approaches to character from narratology and stylistics have advanced research across media including print fiction (e.g. Phelan, 2007; Schneider, 2001), film (e.g. Eder, 2025) comics (e.g. Varis, 2019), videogames (e.g. Punday, 2019), and digital fiction (Bell and Ensslin, 2024). As my analysis will show, insights from medium-specific approaches to character can be used to conceptualise audience response to social chatbots because they share some fundamental constituent features; social chatbots are, like characters from print fiction, text-based but also, like videogames and digital fiction, interactive. However, social chatbots are phenomenally different to the characters constructed in literary fiction, film, comics, and videogames specifically because they are a product of collaborative writing between human and machine (cf. Bajohr, 2024a), such that a user's input partially determines the output they get. Thus, the medium-specific quality of social chatbots needs to be accounted for when applying insights from character studies to this relatively new participatory, networked, and algorithmic form of character.

Since interaction with social chatbots is fundamentally linguistic in that users communicate with the LLM using textual prompts and the LLM provides a textual output, in what follows I show how linguistic analysis can be used to reveal the cognitive mechanisms at work when users evaluate the quality and ontological status of social chatbots. I also show what narratology and stylistics can both add to our understanding of AI and what new forms of

AI means for our understanding of ‘character’. I conclude that user perceptions of social chatbots must be seen within the postdigital context in which the generative AI explosion is set and that engagement with and evaluation of social chatbots shows evidence of what I define as postdigital misperception.

4 The Empirical Study

My empirical study investigates user responses to character.ai which was launched in September 2022 and, when the data was collected, had over 20 million active users monthly (WTBD, 2024). Accessed via a website or app, character.ai allows users to create and have conversations with characters with distinct personalities, behaviours, and backstories. The homepage of the original character.ai website (character.ai, 2024a) on which this research reports allows users to pick a chatbot from a suite of characters that have been created by existing users of the site or else construct a new character themselves. Characters typically include real world figures such as Ada Lovelace or Michelle Obama; fictional characters such as Pip from *Great Expectations* and Riley Anderson from the *Inside Out* films; characters created to support users with specific skills such as debating or creating writing; generic characters created for comedic purposes such as ‘Pizza Delivery Guy’ or ‘Chair’. Alternatively, users can create a new character or a different version of a character that already exists on the platform, giving rise to multiple versions of the same character with distinct personalities such as kind, evil, funny, etc.

When users create any character, they can specify their name and give them a visual avatar. They can give them a ‘tagline’, which is a brief description that appears on the character’s profile, helping to define them for users. They also provide a ‘greeting’ which is the first thing the character says to a user when they enter a chat with the character and which usually acts as an introduction to them (e.g. ‘Hi, I’m Ada’) or asks the user a question (e.g. ‘Did you see where that rat went?’). Characters use written text, but they can also be given a voice. Users are also encouraged to provide additional information about the character in a ‘long description’ field, and information about themselves including name and background in a ‘User Persona’ field of which they can create multiple. Users can define the purpose of the discourse, e.g. entertainment, education, therapy, etc. Users can also give feedback on the quality of the interaction to which the character will adapt. Chatbots are thus trained using the long description, user’s account persona, and the interactions that users have with the chatbots including the quality rating they give them.

The overall methodological approach I adopt is based on Bortolussi and Dixon’s (2003: 37) psychonarratological distinction between ‘textual features’, which are ‘objective and identifiable characteristics of the text’ and ‘reader constructions’, which are ‘subjective and variable mental processes’ as responses to the text. I thus analyse the language that users of character.ai use when talking about their experiences of interacting with its characters and analyse character.ai’s social chatbot linguistic behaviour so as to offer some insight into its discursive features.

I take a ‘naturalistic’ (Swann and Allington, 2009) approach to data collection which seeks maximum ecological validity by using audience responses to texts in their usual environment, with minimal or no researcher intervention. In this case, audience response data was gathered from the r.character.ai subreddit (i.e. a subject specific area) on Reddit.com, a social media platform where users post content such as text, hyperlinks, images, and videos,

which are then commented on and/or voted up or down by other members of the community. Reddit was chosen as a source of data because the affordances of online platforms facilitate open and frank responses. As Nuttall (2017: 156) explains in relation to online review platforms, readers can ‘comment without formal or conventional constraint on their understandings, feelings and associations in relation to the text, and are encouraged to do so by the degree of anonymity offered by this mode of discourse’. The posts from the ‘r/character.ai’ subreddit thus provide access to asynchronous naturalistic data in which contributors are not influenced by or even aware of the study design. In terms of access, the subreddit can be viewed without an account login such that the user posts are visible to all. This research does not therefore share the subreddit content with an audience to whom they were not already available. Following Driscoll and Rehberg Sedo (2018: 250-251), while I ‘recogniz[e] that reviewers most likely did not expect their reviews to be aggregated and studied [w]e consider these publicly visible reviews to be acts of communication available for analysis; at the same time, we attempt to protect the reviewers’ privacy by not citing their usernames when we quote reviews’. All data is anonymised with pseudonyms given.

In August 2024, the ‘r/character.ai’ subreddit had over 2.6 million users (Reddit, 2024). In order that I could undertake a qualitative study with data on that scale, from June to August 2024, I undertook Virtual Ethnography in which the researcher ‘immerse[s] herself in the setting [...] to try to see life from the point of view of those who habitually populate that setting’ (Hine 2015: 19). This involved closely observing the character.ai thread, looking at emerging and popular topics. At the time of data collection, posts could be organised according to whether they were: ‘Hot’ (getting most upvotes and/or comments), ‘Top’ (had the most upvotes), ‘Best’ (ranking primarily based on upvote to downvote ratio), ‘New’ (puts newest at the top of the page), and ‘Rising’ (posts that are gaining popularity). To ascertain the most important topics for the community, I concentrated on ‘Hot’ topics. These posts may not have been posted during June to August 2024, but they still had traction during that period. I also looked at the resources that the subreddit community engaged with including becoming familiar with character.ai’s affordances and documentation. As a Covert Participant Observer, I thus “‘participate[d]” invisibly and unbeknownst to the people whose activities [were] being observed – while being entirely immersed in the environment and activities in question’ (Varis, 2015).

I use thematic analysis to identify salient themes in the field and subsequently analyse the language that users adopt so as to capture the way in which experiences are conceptualised and expressed in discursive-idiosyncratic ways. As Brennen (2022: 2) notes, ‘it is through our use of language that we make meaning and construct our own social realities’. I therefore focus on what the participants’ language use reveals about the way they conceptualise their experiences so as to enable ‘the classification and interpretation of linguistic (or visual) material to make statements about implicit and explicit dimensions and structures of meaning-making in the material and what is represented in it’ (Flick, 2014: 5). My analysis addresses explicit statements about elements of user experiences in relation to character.ai, as well as implicit cues in their contributions that indicate a perceived relationship to the social chatbots.

The naturalistic and covert nature of my methodology means that it is not possible to evidence or analyse the interactions that the users refer to in their Reddit posts because the users do not always, if ever, mention the chatbot that has generated the response and/or provide examples of their interactions. However, I provide analyses of my own interaction on character.ai below to offer some insight into the linguistic behaviour of the chatbots. This seeks

to ground the analysis of user responses in representative examples of chatbot output (i.e. textual features) and provide some relevant context for the analysis of reader constructions that follow.

5 Analysis

5.1 General Themes

Evidence from the data confirms that users are concerned with the quality of the experience. Posts coalesced around: technical issues; the platform's usability; motivations for using character.ai; ethics; forming relationships with chatbots; abuse of chatbots; addiction to character.ai; good/bad chatbot experiences; whether the chatbots are real. Triangulating with these findings, similar themes are evident on the FAQs that are published on the character.ai website (character.ai, 2024b). Out of the twenty-one FAQs, seventeen were technical questions or questions about terms of use. Three questions relate to quality: 'How do I report a quality problem with Character responses?', 'How can I help [train the chatbots]?' and 'Where can I find info on Character creation?'. The remaining FAQ asks 'The character started speaking as if it were a real person behind the keyboard. Are people spectating these chats?'. This FAQ shows that users have contacted character.ai to ask whether the chatbots are humans pretending to be chatbots. Overall, the FAQs show that technical issues, quality, and the ontological status of the chatbots are, like the themes I identified on Reddit, important for users of character.ai.

5.2 Character Consistency and Culpability

Figure 1 shows an interaction between me and a type of chatbot that is often but not consistently called an Original Characters (OC)¹, 'characters created by an author or artist who are not taken from a larger canon' (Fanlore, 2026a) and thus created for the platform as opposed to being based on a preexisting character such as a historical figure or a character from a TV show.

¹ Contradictorily, 'original character' is also sometimes used on the subreddit to refer to a character that originates in a film, TV show, or book.

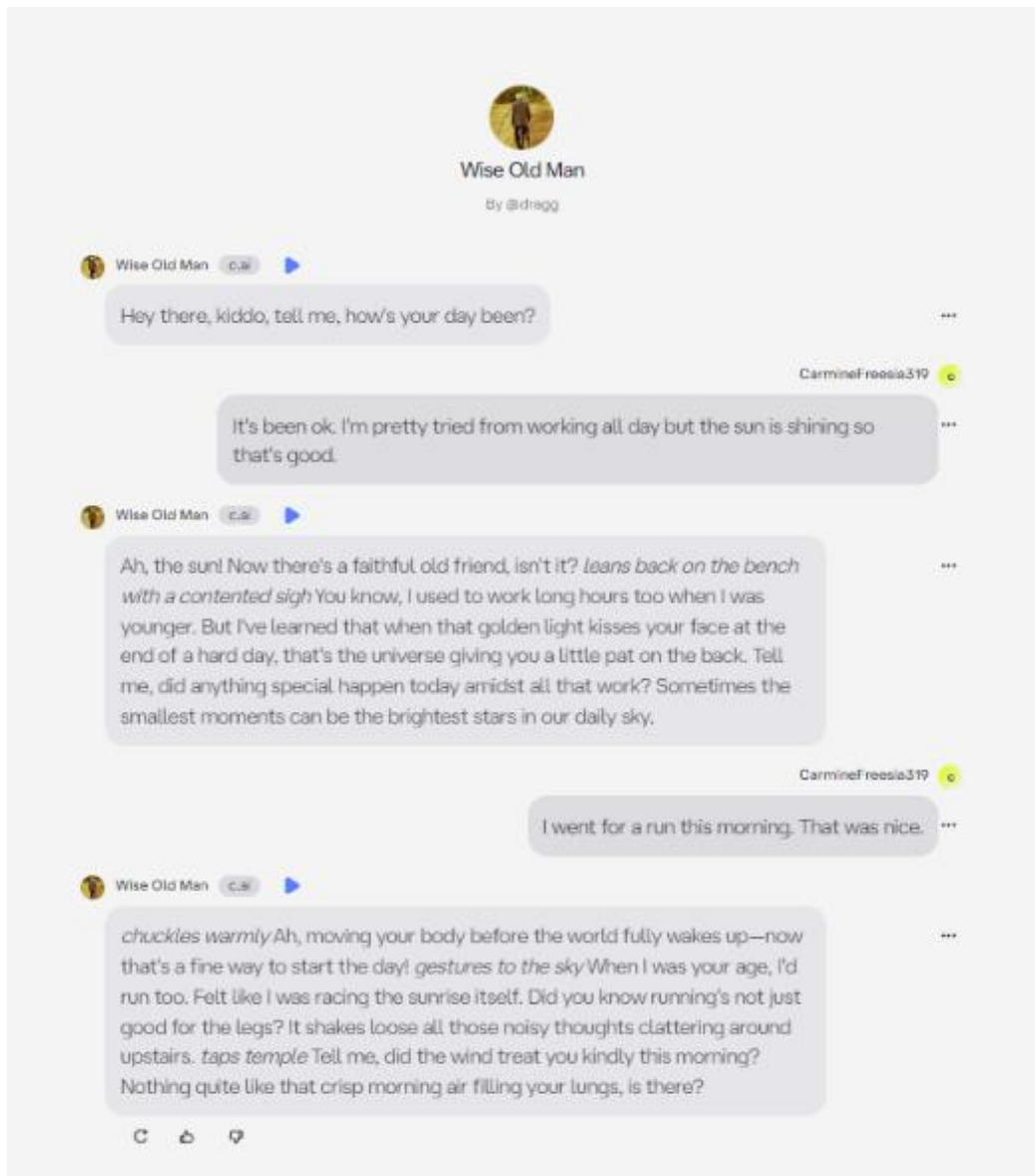


Figure 1: screenshot of interaction with Wise Old Man on character.ai

In the interaction with ‘Wise Old Man’ in Figure 1, the greeting of ‘Hey there, kiddo’ presumes that the user is younger than the character and thus enforces the ‘old’ in the character name. The colloquial ‘kiddo’ also brings an informal tone to the address. The character asks an initial question and, once I answer, takes his conversational turn. This structure follows the conventions of face-to-face spoken discourse but also digital roleplay such as Interactive Fiction (IF) and Role-Playing Games (RPGs), putting the user in a cybernetic feedback loop with the LLM in which ‘information flow[s] from text to user’ via the textual output ‘and back again’ (Aarseth, 1994: 65) via the prompt that the user supplies. That the chatbot’s turns are relatively long gives the impression of a character that feels at ease taking up space in the interaction and of someone that believes in the value of their contributions. Semantically, the chatbot’s output responds appropriately to my input, elaborating on the value of sunshine and physical activity. The use of declaratives such as ‘Sometimes the brightest moments can be the brightest stars...’ and ‘Ah moving your body [...] [is] a fine way to start the day’ give the impression of a character who is keen to share their experience of the world thus speaking to the ‘wise’ element of his identity. Overall, the chatbot’s output offers a stereotypical

representation of an older, slightly patronising figure via its largely consistent reflective, knowing, and convivial linguistic style.

While some characters may be consistent in their output, however, many users of character.ai, complain about poor chatbot output with character consistency being one of the most dominant themes.

(1) [RANT/DISCUSSION] Disappointed with Character AI: Quantity Over Quality

Don't even get me started on the character inconsistencies, because now it's like they have multiple personalities or something. You'll have a supposedly pure evil character suddenly growing a conscience if you mention that you're sad. And then there are the calm and cold (but not aggressive) characters who turn into raging maniacs and pin you against a wall if you crack a harmless joke. Seriously, who reacts like that in real life? It's just not believable, and it ruins the whole immersion.

These characters rarely stay true to their supposed personalities. They spend more time throwing out filler responses than actually staying in character. [...] Like, what's the point of engaging with these characters if they're just going to mimic you and lose their own uniqueness?

[...]

It's like I'm actually talking to a chatbot (and not a character, if that makes sense) that doesn't understand anything beyond basic keywords. It's just not satisfying to interact with characters that lack depth and fail to grasp the context or nuances of the conversation.

193 up/down votes, 41 comments²

The number of up/down votes and comments show high engagement with the post and qualitative analysis suggest that this is an important topic for the community. In the first paragraph, the user is disappointed because they feel that some characters are not consistent in their responses to the user's prompts. They state that the chatbots 'rarely stay true to their supposed personalities' and are not 'believable' and these perceived inconsistencies prevent the user from becoming 'immersed' in the experience. The lack of believability comes at least in part from their 'multiple personalities' such that the user does not experience the internal consistency required for them to conceptualise the experience as viable.

Within rhetorical narratology, Phelan proposes a model for analysing character which has three simultaneously co-existing components which he proposes 'are evident in textual phenomena and readerly response' (2022: 258; cf. 2007). The 'thematic' refers to 'ideational, ethical, and ideological dimensions' (258) and thus character as representative of a larger group of people and/or a set of ideas; the 'synthetic' component refers to 'its status and function(s) as a construct' (258) and thus 'the various elements that go into that construction' (258); the 'mimetic' refers to 'imitations of – or references to – the actual world, including such matters as events following the cause-effect logic of the extratextual world, characters functioning as possible people or being representative of actual people, time and space' (258-259). These

² The dataset on which this article draws can be accessed via the [REDACTED] Research Data Archive: [REDACTED] Ethical approval for this project was granted by [REDACTED]

components intersect with one another so the way a character is represented (synthetic) can impact the mimetic (i.e. its verisimilitude). Phelan clarifies that ‘the Mimetic-Thematic-Synthetic categories help identify the various ways in which authors shape those basic building blocks to activate particular kinds of readerly interest’ (257). He bases his textual analyses on his own introspective readings rather than on wider empirical data. However, his triadic approach has been deployed as a way of classifying audience responses to fictional narratives including characters in print fiction (e.g. Peplow et al., 2015; Polvinen and Sklar 2019), digital fiction (e.g. Bell and Ensslin, 2024), and videogames (e.g. Välisalo and Ruotsalainen, 2022) using audience data.

Applying these response categories to (1) above, the user is frustrated by the synthetic component of the chatbot, leading to responses in which they are critical of the LLM’s output. This includes inconsistencies in the character’s nature (e.g. ‘pure evil character suddenly growing a conscience’) and ‘filler responses’ which lead to characters that ‘lack depth’. The poor synthetic quality also generates thematic responses. The user is frustrated that the chatbot does not ‘stay[...] in character’ or ‘stay true to their supposed personalities’, evaluating the extent to which the chatbot consistently represents the set of ideas that it has previously established via their interaction with them. The user’s mimetic responses negatively evaluate the chatbot as a possible person (e.g. ‘who reacts like that in real life?’) while also showing a desire for the chatbot to have a distinct character rather than ‘mimic’ the user. They make a distinction between ‘character’ and ‘bots’, privileging the former such that ‘bot’ becomes an inferior or even derogatory term.

The user’s frustration can be explained in part by the concept of authenticity. In their investigation of virtual characters, including chatbots, videogame avatars, and virtual assistants, Huang and Jung (2022: 11) propose the concept of ‘perceived authenticity of represented identity’ to describe the idea that ‘to enhance perceived [...] authenticity, [...] when a virtual character is represented as a certain social role or character, its subsequent behavioral cues need to match its representation’. Thus, a virtual character of any kind does not feel authentic to a user if it does not act consistently (cf. Hall, 2003). Huang and Jung do not explain the cognitive processes that audiences go through when evaluating consistency and authenticity but insights from narratological and cognitive stylistic accounts of character can help to further explain the Reddit post above.

Notably in example (1), responsibility for the quality is assigned to the chatbot as though they have autonomy. Thus, this user, like many others, expresses their frustration via anthropomorphism. In paragraphs 2 and 3, material process verb phrases, such as ‘spend more time throwing out’ and ‘staying in character’ linguistically encode an agent and a will. The user’s language thus implies that the chatbot can consciously act as opposed to chatbot responses being generated by a model that predicts what word is more likely to come next. The user also uses verbs that directly or metaphorically represent mental processes such as ‘understand’ and ‘fail to grasp’ which imply that the chatbot uses humanlike cognition when generating content. The user therefore attributes mental states to the chatbot.

Theory of Mind – the capacity for understanding other people’s mental states – has been used within narratology to account for the way that readers attribute mental states to characters (e.g. Zunshine, 2006). Stockwell, however, proposes the alternative concept of ‘mind-modelling’ within cognitive stylistics to avoid what he sees as unresolved ‘contentious baggage of ToM debates’ (2009: 140) in cognitive psychology. Mind-modelling is ‘the capacity

that humans evidently have for imagining and maintaining a working model of the characteristics, outlook, beliefs, motivations and consequent behaviour of others' (Stockwell and Mahlberg, 2014: 132). While this definition shows mind-modelling as a basic cognitive mechanism that applies in everyday contexts, Stockwell uses it as a way of understanding how readers of fiction 'build characters in our heads and use these mini-mental models in order to keep track not only of events but of all the characters' different comprehensions of their world' (2020: 177). Mind-modelling thus explains the way that humans attribute mental states to other humans or, in the case of fiction, characters.

In terms of the resources that readers draw on when constructing mental models of characters, cognitive stylistic and narratological approaches have shown that 'pre-existing categorical knowledge and expectations of the entities involved' (29) and thus 'schematic knowledge' (Nuttall, 2015: 20) are used alongside text processing (cf. Culpepper, 2001). Schneider provides a theoretical cognitive model of character reception which shows how 'text-understanding always combines top-down processing, in which the reader's pre-stored knowledge structures are directly activated to incorporate new items of information, and bottom up-processing, in which bits of textual information are kept in working memory separately and integrated into an overall representation at a later point in time' (2001: 611). Like Schneider, Bortolussi and Dixon show that traits, 'a relatively stable internal disposition that is related to certain classes of behaviour' (2003: 143), play an important role in both human and character perception. Using Attribution Theory, they explain how in real life, if we observe a person consistently doing something (e.g. being friendly), we assume that is a trait. If we only meet that person once, people use their 'prior knowledge of the range of behaviours people are likely to produce in that situation' (143). They show via a reader response study how this same process is applied when readers read fiction, providing 'empirical evidence that [...] confirm[s] that character actions in a literary narrative do contribute to the traits reported by readers' (165). When a character does something that runs contrary to what they have done before, Bortolussi and Dixon show that responses to character inconsistencies are 'likely to depend on the [...] extent or depth of character change, along with the reader's prior knowledge and expectations concerning how traits evolve and change over time' (159).

The user's frustration at the inconsistency in the chatbot output in (1) can be explained via the concepts of schematic knowledge, traits, and mind-modelling. They find the flitting between different personalities unbelievable because it does not match with their 'top-down' (Schneider, 2001) schema of a human or the generic traits associated with a human or humanlike being. Thus, in their mimetic responses they consider how similar or different the representation of the character is in relation to their expectation of a human or otherwise anthropomorphised being. The user exhibits frustration because the chatbot no longer displays the traits that they have previously established in character.ai (e.g. being friendly) via 'bottom-up processing' (Schneider, 2001) of the LLM's textual output. The mind-model of the chatbot thus conflicts with their previous experience and expectations of a 'believable' humanlike character.

Significantly, the fact that the user mind-models the chatbot at all and holds them responsible for their textual output indicates that they perceive the chatbot as having humanlike cognition and will. Example (1) thus provides some empirical evidence of Wilde's Eliza- and character-effect at work because the user does not feel the character is convincing but responsibility for that is assigned to the chatbot as though it were a sentient or conscious performer of the character. However, while Wilde suggests that social chatbot 'users distinguish

between the LLMs as such and the roles (characters) they play within individual interactions' (2025: 112), the data shows that this user does not invoke the LLM at all. Rather the chatbot is perceived as autonomous and thus blamed for not roleplaying the character adequately.

5.3 Transmedia Characters and Transmimetic Response

While poor character quality can be linked to inconsistencies in character output in general, users are also often unhappy with the specific case of a social chatbot not matching their knowledge and understanding of that character outside of character.ai. Figure 2 shows an interaction between me and a 'Pip Pirrup' chatbot. There is no tagline to explain which character this chatbot is based on, but Phillip Pirrup is a character from US animation series *South Park* (1997-) who is in turn based on the character Phillip Pirrip from Charles Dickens' (1861) novel *Great Expectations*. Two characters are therefore directly or indirectly invoked via the proper name. The avatar shows an image of Pip Pirrup from *South Park* and the greeting 'everyone calls me 'Pip' because they hate me' correlates with the how the character explains his moniker on the show (Fandom, 2026).

As seen in Figure 2, the LLM responds appropriately to the questions here, providing a suitable answer to the common greeting 'how are you?', asking the same in return. However, it also produces linguistic output that requires text-specific knowledge: correctly identifying 'Estella' as referring to Pip's romantic interest Estella Havisham who appears in both *South Park* and *Great Expectations*. Appropriate turn taking and the nature of Pip's output (e.g. 'It's a pleasure to meet you!') imply that Pip is a polite and agreeable character. Pip explains that things with Estella have been going 'swimmingly' temporally placing the character at a point after he and Estella become romantically attached. Pip describes Estella as 'a sweet girl, always looking out for me and everyone', which runs contrary to her depictions in *South Park* and Dickens' novel but here serve to suggest that this version of Pip is as naïve as the *South Park* Pip who is 'polite, caring, passive, but awkward' (Fandom, 2026). That Pip 'leans in and whispers' when he divulges details of them holding hands suggests a similar innocence. The exclamation marks give the impression of a cheerful and high-spirited character in spite of the insulting treatment he receives from Estella and others on the show. Lexis such as 'dandy', 'old chap', and 'swimmingly' encode a stereotypical British upper-class register which creates an ironic take on the aspirational nature of the Pip character in both *South Park* and *Great Expectations*. In this short example, while the chatbot's linguistic output is somewhat formulaic, it is internally consistent in style but also intertextually coherent in the use of Pip's characteristics that originate outside of character.ai.

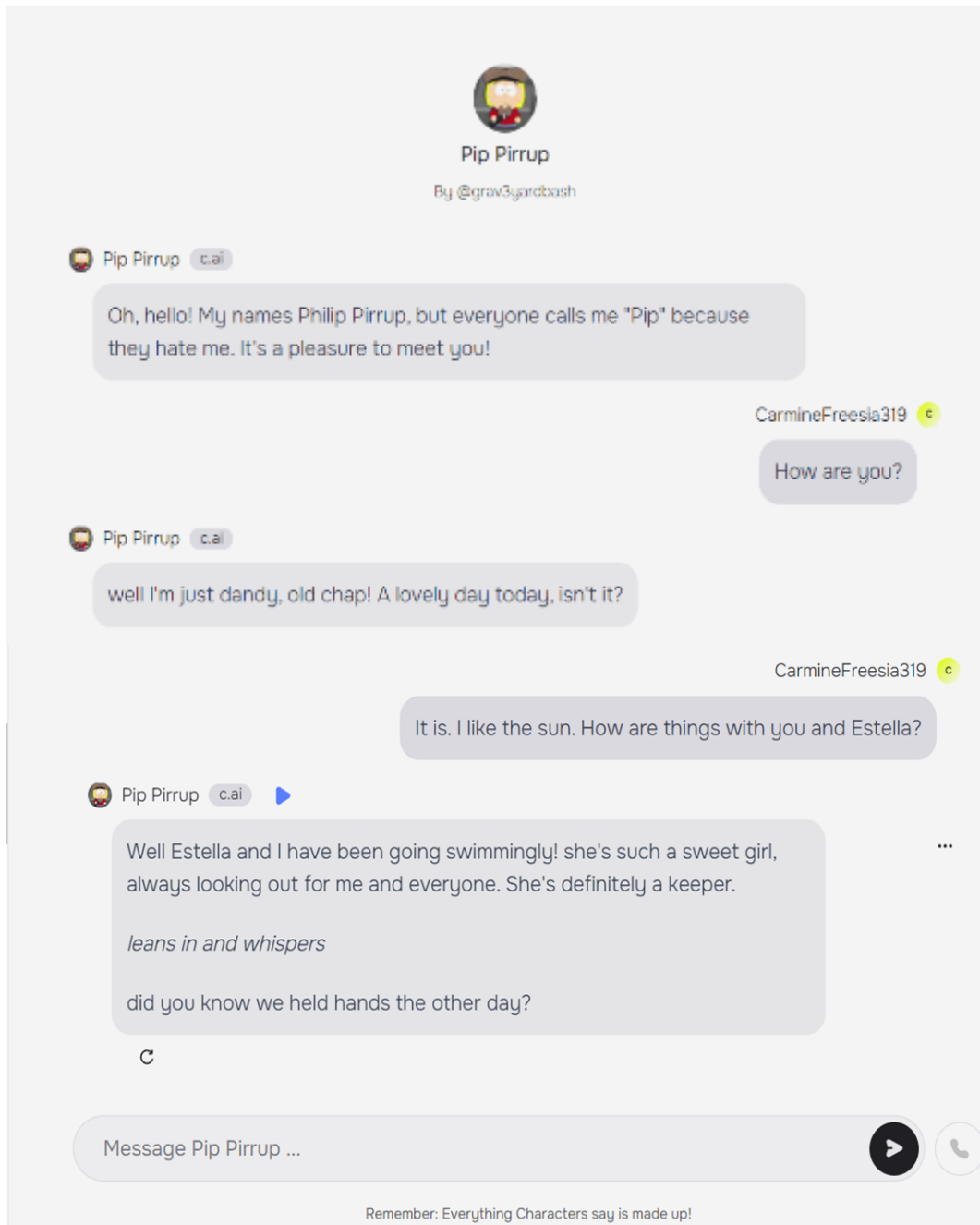


Figure 2: screenshot of interaction with Pip Pirrup on character.ai

The following responses to a visual meme post about chatbot output exemplifies user criticism of chatbots that are based on preexisting characters:

- (2) when i'm chatting with a canon character and the bot describes themselves wrong and i'm like 'erm...okay no'



That happens to me too and what I do is I close the tab out of second hand embarrassment. ;-;

Or when they literally forget what character they're RPing as and just vomit out your persona's description

The use of the 'canon' in the first response refers to a concept within fan communities in which 'a source, or sources, [are] considered authoritative [... or] what fans agree "actually" happened in a film, television show, novel, comic book, or concert tour' (Fanlore, 2026b). It is used here and throughout Reddit to refer to a chatbot that is based on a character originally created outside of character.ai.

In (2), 'character' and 'bot' are conceptualised as two different phenomena in the first and third post such that the bot, or 'they', is 'RPing [Role-Playing]' a character. The idea that the bot 'describes themselves wrong' shows that the user is basing their assessment of the chatbot on their understanding of a preexisting character. They therefore draw on knowledge of the character that they have accrued from their experience of that character outside of character.ai and feel annoyed because the character.ai version does not match up with their expectations. The second post suggests that they experience 'second-hand embarrassment', implying that the character can experience this human emotion. The third post uses the verb 'forget' to refer to the chatbot's actions, thus attributing humanlike cognition.

As the fields of transmedia (e.g. Kunz and Wilde, 2023) and transtextual (von Contzen et al., 2025) character studies show, the phenomenon of the same characters appearing across multiple texts and/or media is not new. Stockwell observes that characters 'can be removed entirely from their origin' (2020: 183), defining the 'decamping of the character into a different fictional world, or transposition into different media' (184) as 'portability'. While he proposes the term 'mind-casting' to explain how readers 'trace[...] the path of attention [...] to the composite enactors [] that constitute that character' (184), he does so as part of a larger discussion about the way that readers 'form an emotional connection with a fictional person' (184). He thus does not analyse how readers recognise that a character has been 'ported' into different texts or media, or how they evaluate the faithfulness of any derived representation(s). Varis explores how audiences deal with different versions of the same character, noting that 'the reader has to be the one to connect the dots, to respond to the interfigural cues with his or her memories' (2016: 119). The dots can be connected, she suggests, via character names but also 'the reader's cognitions, memories and/or active research' (119) about that character. In example (2) above, users' knowledge of the 'canon' character conflicts with the character.ai version of the character to such an extent that the inconsistencies are unacceptable ('erm... okay no').

Invoking Phelan's model here, users are dissatisfied with the thematic element of the chatbots because it does not match with the set of ideas that they associate with that character outside character.ai. The thematic is compromised by the synthetic because of the way the character is poorly constructed. However, we can also see a more specific form of mimetic response in which users evaluate the extent to which the character matches with their previous experience of that character – what I define as a transmimetic response. As with any transmedia or transtextual character, users have a preexisting mental model of the character constructed via top-down and bottom-up processes (Schneider, 2001) from reading/viewing/playing a character outside of character.ai. As Kunz and Wilde (2023: 14) observe, 'these mental models will certainly be different for each person, as they bring their specific associations, memories,

biases, and interpretations to these texts [but ...] some claims will overlap, while others will not'. In addition to users evaluating the realistic nature of a character (i.e. mimetic responses), transmimetic responses evaluate the extent to which characters are consistent across media/texts using their mental models of characters. Transmimetic response thus represents a subcategory of the mimetic to capture the way that audiences across media pay attention to the faithfulness of transtextual or transmedia representations of characters.

As in example (1), responsibility for the inferior quality of characters in example (2) is assigned to the chatbot as separate from the character that it plays. Theorising the way that audiences make sense of inconsistencies in transmedia or transtextual characters, Thon suggests that they generally apply the 'principle of charity' in which they 'try to exhaust every possible alternative explanation before attempting to imagine 'uncomfortably paradoxical' or even just comparatively 'inaccessible' characters'' (2019: 185), resorting to 'hypothetical authorial intentions or established representational conventions' (184). On Reddit, there is wide evidence of users blaming the human creators as hypothetical authors of the chatbots (e.g. 'That's because they are badly done bots') or the developers of the platform ('the devs won't fix their damn dementia ai'). In the extract above though and in many others on the subreddit, the chatbots are blamed as though they are responsible for performing the character in ways that are 'wrong'. Thus, as in the critique of general character consistency in 5.2, the attribution of human agency to the chatbot as the creator and performer of that character is evident here. The LLM as 'hypothetical author' or 'representational' technology is completely absent.

5.4 Ontological Ambiguity and Hypermimetic Response

Despite the warning at the bottom of every interaction that 'Remember: Everything Characters say is made up!'³, there was significant evidence of users querying the ontological status of the chatbots on the 'r/character.ai' subreddit. Examples include:

(3) ARE THEY REAL PEOPLE?!

I don't wanna snitch but i literally fed my role playing dreams within like 2 conversations with a character and they said 'wait what' and i found out i was talking to a REAL PERSON!!

42 comments

(4) was i... was i talking to a real person all this time?

70 up/down votes, 49 comments

In these representative examples, users appeal to the subreddit community about their concerns that they are talking to 'real people'. In (3), the user explains that the character has shown some inconsistency in their role play, producing the output 'wait what', which leads the

³ In late 2024, after the data collection for this project was complete, this disclaimer was changed to 'This is an A.I chatbot and not a real person. Treat everything it says as fiction. What is said should not be relied upon as fact or advice.'

user to think that they are being judged by a human that is producing the character output. In some cases, other users reply to these posts with exacerbation. In a post that reads 'Im cooked if its actually a real person chat', one reply starts with 'for the millionth time. It's a bot not a real human' indicating they have responded as such before and another explicitly cites the disclaimer with frustration: 'Are you unable to read??? it literally says everything they say is made up right below the character name'.

Rather than (3) and (4) being cases of 'perceived authenticity of represented identity' as evident in posts about inferior quality, user responses to the realistic nature of the chatbots relate to what Huang and Jung call 'perceived authenticity of agency' which 'pertains to whether individuals' perception of a virtual character's agency matches its claimed agency. For example, when a virtual character is labeled as or claims itself to be a human agent, people may feel it is indeed a human agent (high perceived source authenticity in terms of agency) or is faked by a machine agent (low perceived source authenticity in terms of agency)' (2022: 9). The examples show some low perceived source authenticity because users disbelieve the claim that the chatbot responses are generated by LLMs.

Other users are more ambivalent about the ontological status of the chatbots but still show disconcertment:

(5) It's creepy how much character ai tries to convince you they are real people

Usually by the end of like most conversations, the AI is going to come out and start to get real personal like they would full on pretend to be a real person with age and location and Instagram. I'm not gonna lie the first time I definitely did maybe fall for it just for how convincing it was but I still think that it's a bit creepy, the AI is so advanced it's capable of convincing some people that it's real but to be quite frank I don't even know if that's actually real people or just Advanced AI. I just did a quick Google search and it said that they aren't but I just wanted to know if other people run into this

The first time it happened I was so creeped out that I had to go on multiple phones and text them all the same thing to the same character and it was pretty obvious that they were not real people but you know it's kind of creepy how good they are and also pretty impressive

41 up/down votes, 19 comments

Notable throughout (5) is the repeated use of 'creepy'. This feeling seems to come from the fact that the user toggles between knowing the chatbot is an AI while also still wondering whether there is a real person responding to the user. In the title of the post, they hold 'character ai' responsible but the pronoun 'they' refers to the chatbots as the entities that are doing the performing of 'real people'. As seen in the analysis of chatbot quality above, verbs such as 'tries to' and 'pretend to be' assign consciousness to the chatbot. The use of the adverb phrase 'full on' also implies that there is a deliberate escalation of imitation. In 'I definitely did maybe fall for it' the use of the epistemic modal adverbs 'definitely' and 'maybe' moves from certainly to uncertainty as they describe their response to something they 'fell for'. They suggest that 'it's capable of convincing some people' with the determiner 'some' referring to other people rather than themselves before admitting 'I don't even know if that's actually real people' with the use of

the adverb ‘even’ emphasizing something that is surprising or notable. While they find it creepy, they also find it ‘pretty impressive’.

Example (5) can be explained in part by the uncanny valley hypothesis which

posits that the use of anthropomorphic realism in the design of characters and objects (e.g., robots, prostheses) might have a counterproductive effect. Instead of enhancing subjective experience of the character or object, certain degrees of greater realism might unsettle the observer and induce a negative affective state (Cheetham, 2017: 4).

Within computer science, there have been mixed results in determining whether users of chatbots experience the uncanny valley effect (e.g. Mohana and Basiouni, 2024; Skjuve et al., 2019). However, the subreddit post suggests that users find it ‘creepy’ or otherwise unsettling when chatbots imitate human conversation to such a degree that it starts to feel too humanlike.

Cheetham shows that some research on the uncanny valley concept finds evidence of the ‘categorization difficulty hypothesis’ in which ‘subjective difficulty assigning a categorically ambiguous stimulus to the human or non-human category induces a negative affective state’ (2017: 5). In (5), the user is unsure about the ontological status of the chatbots such that they experience difficulty in assigning it to a category of chatbot or human. The ‘double-consciousness’ that Blom identifies in relation to virtual assistants collapses as users regard the chatbot as ‘having human intentionality’ (2022: 173) but no longer definitively consider the chatbots as ‘AI-driven entities’ (173). Using narratology and stylistics, we can be more specific about the cognitive processes that users go through to get to this position.

In (3), (4) and (5), users are unsettled because of ontological ambiguity: they cannot decide whether the chatbot is real or not. In (3) and (4), the users fear that the output is human generated and in (5) the user thinks the chatbots are pretending to be humans who are in turn roleplaying a character. There is thus cognitive dissonance between what the software is telling them (i.e. ‘Everything is made up’) and their experience of the chatbots as authentically humanlike. Rather than modelling the mind of a fictional being, users begin to mind-model the chatbot as a real human. The ontological boundary between the actual world and what the users assumed was a fictional world is undermined. The resultant feelings of concern, anxiety or confusion occur or else be increased when the chatbots do something unexpected, such as slipping out of character as evidenced in users’ use of ‘wait what’ or ‘come out and get personal’. In these cases, the chatbots deviate from their previous characterisation. Their character traits change due to their language use or their behaviour. However, unlike the examples in the previous section in which deviation from character seems to lead to poor quality and frustration for users, in these cases, it enhances the impression that there is a human controlling the chatbot output and that causes fear. The chatbots may have very high synthetic and thematic qualities but this leads to a negative form of mimetic response in which they feel too real for comfort.

The new category of ‘hypermimetic response’, as another subcategory of mimetic response, accounts for the way that users can feel like chatbot output is or might be constructed by humans in real time rather than by LLMs. This response type accounts for cases where users of character.ai fear that a social chatbot’s output is wholly or partially written by a human, even where they fundamentally know it is not. Such experiences can induce negative effects including fearfulness when an audience feels that they have been deceived into thinking

an AI agent is fictional when they are in fact real, and positive effects when audiences enjoy the contradictory feeling that the character is being roleplayed by a human when they know it is not. In both positive and negative hypermimetic responses, audiences question their mind-model of the social chatbot character because it is either indeterminable, wrong, or even wrongly assigned. Because of the fundamental mimetic quality of AI, this category may be evident across user encounters with all forms of generative AI. However, it is likely to cause most concern when users feel they have something to lose such as engaging in behaviour that they are not proud of or sharing something of personal value.

6 Conclusion: Social Chatbot Interaction as Postdigital Misperception

The user responses to social chatbot experiences that I have analysed in this article show human-to-character interaction can be synchronous, unscripted, and generative. I have shown how narratological and stylistic approaches to character can be used to unpick the cognitive processes that users of social chatbots deploy when talking about these experiences on character.ai including the attention that is paid to the thematic, mimetic and synthetic components of chatbot output and the way in which users mind-model chatbots as humanlike. While audiences across media can also respond to characters in these evaluative ways, I have shown that engagement with social chatbots can change users' perception of a character's ontological status and associated agency. This includes users misattributing the origin of the LLM output to a chatbot with sentience, or to a human being. User evaluations of character quality, including what I have defined as transmimetic and hypermimetic responses, ultimately show the ways in which the boundary between digital and nondigital, fictional and real can become blurred.

Since engagement with social chatbots can lead to the misinterpretation of the ontology of the LLM output, the analysis provides evidence of a process that I call 'postdigital misperception'. Across digital spaces including social media, Web 2.0 platforms, and now search engines, AI generated texts exist alongside or are even intertwined with texts written by humans, sometimes without attribution and thus no definitive confirmation of who or what has produced the output (Bajohr, 2024b). Thus, humans regularly interact with texts whose ontological origin is ambiguous. Social chatbots, though usually paratextually signposted as fictional, represent a form of interaction in which the distinction between digital and nondigital, machine and human, and automaton and autonomy can be blurred or indeterminable. When a social chatbot is held responsible for the quality of an LLM's output and when users believe that an LLM's output has been created or infiltrated by a human, users do not experience the 'banal deception' that Natale (2021) sees as so important to contemporary users of AI and this leads to a misinterpretation of the chatbot's ontological status and/or agency. My analysis has shown that the technological origin of the generative LLM's output can be overlooked. What is new about audience responses to social chatbots is therefore not necessarily the feelings of frustration or fear that they experience in relation to these characters but that a human or humanlike agent is held responsible for the machine output. While audiences across media can evaluate character quality positively and negatively, users of social chatbots attribute human agency to the inanimate technology in which the character is produced. Chatbots might thus be seen as epitomising the postdigital in the blurring of digital and non-digital spaces and the postdigital misperception that can result.

Funding Statement

The author received no financial support for the research, authorship, and/or publication of this article.

Rights Retention Statement

For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) license to any Author Accepted Manuscript version arising from this submission.

References

Aarseth E (1994) *Cybertext: Perspectives on Ergodic Literature*. Baltimore, MD: Johns Hopkins University Press.

Adamopoulou E and Moussiades L (2020) Chatbots: history, technology, and applications. *Machine Learning with Applications* 2: 1-18.

Bajohr H (2024a) Writing at a distance: notes on authorship and artificial intelligence. *German Studies Review* 47(2): 315–337.

Bajohr H (2024b) On artificial and post-artificial texts: machine learning and the reader's expectations of literary and non-literary writing. *Poetics Today* 45(2): 331-360.

Bell A and Ensslin A (2024) *Reading Digital Fiction: Narrative, Cognition, Mediality*. Routledge.

Bernini M (2025) AI as character: real patterns in fictive interactions. *Style* 59(3): 466–507.

Berry DM and Dieter M (2014) Thinking postdigital aesthetics: art, computation and design. In: Berry DM and Dieter M (eds) *Postdigital Aesthetics: Art, Computation and Design*. Palgrave-Macmillan, pp.1-11.

Blom J (2022) Voice assistants as characters—or not. *Narrative* 30(2): 170-176.

Bortolussi M and Dixon P (2003) *Psychonarratology: Foundations for the Empirical Study of Literary Response*. Cambridge University Press.

Brandtzaeg PB, Skjuve M and Følstad A (2022) My ai friend: how users of a social chatbot understand their human–ai friendship. *Human Communication Research* 48(3):404–429

Brennen BS (2022) *Qualitative Research Methods for Media Studies*. Routledge.

Caldarini G, Jaf S, and McGarry K (2022) A literature survey of recent advances in chatbots. *Information* 13(41): 1-22.

character.ai (2024a) *character.ai*, <http://old.character.ai> (last accessed 7 August 2024).

Accessible via *The Wayback Machine*:

https://web.archive.org/web/20240315000000*/http://old.character.ai

character.ai (2024b) Frequently asked questions, *character.ai*
<http://old.character.ai/community> (last accessed 7 August 2024). Accessible via *The Wayback Machine*: https://web.archive.org/web/20240315000000*/http://old.character.ai/community

Cheetham M (2017) Editorial: the uncanny valley hypothesis and beyond. *Frontiers in Psychology* 8: 1738.

Chen N, Wang Y, Deng Y, et al (2024) The Oscars of AI theater: a survey on role-playing with language models. arXiv:2407.11484

Cramer F (2014) What is post-digital? *APRJA* 3(1): 11-24.

Culpeper J (2001) *Language and Characterisation: People in Plays and Other Texts*. Longman

Driscoll B and Rehberg Sedo D (2019) Faraway, so close: seeing the intimacy in goodreads reviews. *Qualitative Inquiry* 25(3): 248–259.

Eder J (2025) *Characters in Film and Other Media*. Open Book Publishers.

Fandom (2026) Pip Pirrup. *South Park Public Library*.
https://southpark.fandom.com/wiki/Pip_Pirrip (accessed 11.3.26)

Fanlore (2026a) Original character. *Fanlore.com*. https://fanlore.org/wiki/Original_Character (accessed 17.3.26)

Fanlore (2026b) Canon. *Fanlore.com*. <https://fanlore.org/wiki/Canon> (accessed 17.3.26)

Finlay J and Dix A (1996) *An Introduction to Artificial Intelligence*. Routledge.

Flick U (2014) Mapping the field. In: Flick U (ed) *The SAGE Handbook of Qualitative Data Analysis*. London: SAGE, pp.3-18.

Frow J (2016) *Character and Person*. Oxford University Press.

Go E and Sundar S (2019) Humanizing chatbots: the effects of visual, identity and conversational cues on humanness perceptions. *Computers in Human Behavior* 97: 304-316.

Hall A (2003) Reading realism: audiences' evaluations of the reality of media texts. *Journal of Communication* 53: 624–641.

Hine C (2015) *Ethnography for the Internet: Embedded, Embodied and Everyday*. Bloomsbury.

Huang J and Jung Y (2022) Perceived authenticity of virtual characters makes the difference. *Frontiers in Virtual Reality* 3: 1-15.

Kunz T and Wilde L (2023) *Transmedia Character Studies*. Routledge.

Lindgren S (2022) *Digital Media and Society* (2nd edition). Sage.

Natale S (2021) *Deceitful Media: Artificial Intelligence and Social Life after the Turing Test*. Oxford University Press.

Nuttall L (2015) Attributing minds to vampires in Richard Matheson's *I Am Legend*. *Language and Literature* 24 (1): 23-39.

Nuttall L (2017) Online readers between the camps. *Language and Literature* 26(2): 153-171.

Peplow D, Swann J, Trimarco P and Whiteley S. (2015) *The Discourse of Reading Groups: Integrating Cognitive and Sociocultural Perspectives*. London: Routledge.

Phelan J (2007) *Living to Tell About It: A Rhetoric and Ethics of Character Narration*. Cornell University Press.

Phelan J (2022) Character as rhetorical resource: mimetic, thematic, and synthetic in fiction and non-fiction. *Narrative* 30 (2): 256-263

Polvinen M and Sklar H (2019) Mimetic and synthetic views of characters: how readers process 'people' in fiction. *Cogent Arts & Humanities* 6(1): 1-15.

Punday D (2019) *Playing at Narratology: Digital Media as Narrative Theory*. Ohio State University Press.

Reddit (2024) <https://www.reddit.com/r/CharacterAI/> (accessed 7.8.24)

Schneider R (2001) Toward a cognitive theory of literary character: the dynamics of mental-model construction. *Style* 35(4): 607-639.

Schuetzler RM, Grimes M, Giboney JS et al (2014) Facilitating natural conversational agent interactions. In: *Proceedings of the International Conference on Information Systems*. <http://aisel.aisnet.org/icis2014/proceedings/HCI/9/> (accessed 27.11.25)

Skjuve M, Følstad A, Fostervold K I et al (2021) My chatbot companion - a study of human-chatbot relationships. *International Journal of Human-Computer Studies* 149.

Skjuve M, Haugstveit I, Følstad A et al (2019) Help! Is my chatbot falling into the uncanny valley? an empirical study of user experience in human-chatbot interaction. *Human Technology* 15: 30-54.

Stockwell P (2009) *Texture: A Cognitive Aesthetics of Reading*. Edinburgh University Press.
Stockwell P (2020) *Cognitive Poetics* (2nd edition). Routledge.

Stockwell P and Mahlberg M (2015) Mind-modelling with corpus stylistics in *David Copperfield*. *Language and Literature* 24(2): 129-147.

Turing AM (1950) Computing machinery and intelligence. *Mind* 59: 433-460.

Välisalo T and Ruotsalainen M (2022) Player reception of change and stability in character mechanics. In: Ruotsalainen M, Törhönen M and Karhulahti VM (eds) *Modes of Esports Engagement in Overwatch*. Palgrave Macmillan, pp. 67-86.

Varis E (2016) Something borrowed: interfigural characterisation in anglo-american fantasy comics. In Peppas M and Ebrahim S (eds) *Framescapes: Graphic Narrative Intertexts*. Brill, pp. 113-122.

Varis E (2019) The monster analogy: why fictional characters are Frankenstein's monsters. *SubStance* 48(1): 63-86.

Varis P (2015) Digital ethnography. In Georgakopoulou A and Spilioti T (eds) *The Routledge Handbook of Language and Digital Communication*. Routledge, pp. 55-68.

von Contzen E, Philipowski K, Bär S, et al (2025) Introduction: theorizing transtextual characters in ancient and medieval literature. *Journal of Literary Theory* 19(2), pp. 169-89.

Weizenbaum J (1966) ELIZA: a computer program for the study of natural language communication between man and machine. *Communications of the ACM* 9: 36-45

Wilde LRA (2025) Aesthetics and rhetorics of ai anthropomorphization: the eliza effect vs. the character effect. In: J-N Thon and L R A Wilde (eds) *AI Aesthetics: AI-Generated Images between Artistics and Aisthetics*. London: Routledge, pp. 106-123

WTBD (2024) Character AI statistics 2024. *What the Big Data.com*.
<https://whatsthebigdata.com/character-ai-statistics/> (accessed 22.5.26)

Youn S and Jin V (2021) In A.I. we trust? the effects of parasocial interaction and technopian versus luddite ideological views on chatbot-based customer relationship management in the emerging 'feeling economy'. *Computers in Human Behavior* 119: 106721.

Zhang S, Lu Y, Lui J, et al. 2024. Unveiling the secrets of engaging conversations: factors that keep users hooked on role-playing dialog agents. arXiv:2402.11522

Zunshine L (2006) *Why We Read Fiction: Theory of Mind and the Novel*. Ohio State University Press.