AI & SOCIETY (2026) 41:3649–3666
https://doi.org/10.1007/s00146-025-02661-8

OPEN FORUM

Undoing Gracia: queering the self in the algorithmic borderlands

Grace Leonora Turtle[1] · **Elisa Giaccardi**[2] · **Roy Bendor**[1]

Received: 14 April 2025 / Accepted: 24 September 2025 / Published online: 4 December 2025

Abstract
This paper introduces *queering as a methodological intervention* in human–AI entanglements, aimed at pluralizing the self and resisting normative algorithmic logics. Expanding the role of queerness in HCI beyond its traditional associations with gender and sexuality, we conceptualize queering as a *relational, performative, and political disturbance* that disrupts the interpellative forces of AI systems. Our exploration unfolds through *Undoing Gracia*, an autotheoretical experiment seeded with the first author's autobiographical memories and values, in which Grace interacts with two digital twin agents, Lex and Tortugi, within the speculative world of Gracia. This multi-agent simulation probes the algorithmic borderlands of subjectivity, as the self is iteratively co-performed and transformed through interaction with the agents. Rather than focusing on technical intervention, the experiment explores co-performance, opening new directions for designing human–AI relations grounded in relationality, plurality, and speculative experimentation. The first author designed and performed the autotheoretical experiment, while the co-authors contributed to the theoretical articulation and critical analysis.

Keywords Queering · Algorithmic borderlands · Multi-agent simulation · Digital twins · Autotheory

1 Introduction

This paper challenges fixed or binary orientations of the self within human–AI entanglements by introducing *queering* as a methodological intervention to pluralize notions of subjectivity entangled with algorithmic systems. It does so by drawing on speculative and situated entanglements between human and AI agents, revealing subjectivity as distributed, relational, and plural. The pluralization of subjectivity becomes especially urgent as a response to the ways in which the rapid rise of datafication and algorithmization has fundamentally reshaped the dynamics of human experience, transforming individuals into fixed "data subjects" locked into normed sociotechnical processes. Through datafication, presumed calculable characteristics of subjects—from sex, gender, dis/ability, race, and ethnicity to geographical markers such as street addresses—are treated as "facts" and used to predict behavior and derive insights (Hong 2020: 11). For instance, predictive policing algorithms point to how data-driven systems risk constraining human agency by analyzing historical crime data which is then used to predict and determine police deployment, and identify and target perceived as "high-risk" individuals. This then reinforces systemic biases that lead to the over-policing of certain communities (Brayne 2020: 5). As Luciana (Parisi (2019): 33) argues, these processes of datafication steer self-determining subjects "towards certain actions so as to benefit the system upon which they depend". In short, being present as data in algorithmic infrastructures, or undergoing processes of becoming data subjects, conditions how one may act within and outside computational worlds, rendering the self increasingly legible to machines while simultaneously curtailing one's capacity for autonomy and transformation.

Parallel to this, research in human–computer interaction (HCI) has evolved from human-centric paradigms to embrace more-than-human, relational approaches (Forlano 2017: 17; Giaccardi et al. 2024: 36-38; Wakkary 2021): 135). These developments emphasize subjectivity as fluid and dynamic, shifting away from static conceptions of the self. Emerging first-person and speculative methods in HCI and design (Desjardins et al. 2021: 3-5; Kinnee et al. 2023:

✉ Grace Leonora Turtle
g.l.turtle@tudelft.nl

✉ Elisa Giaccardi
elisa.giaccardi@polimi.it

✉ Roy Bendor
r.bendor@tudelft.nl

[1] Delft University of Technology, Delft, The Netherlands

[2] Politecnico di Milano, Milan, Italy

4-5; Lucero et al., 2019), such as autoethnography (Ellis et al. 2011: 3-5) and autotheoretical experimentation (Turtle 2022: 3-4), have opened up avenues for exploring entangled human–AI relations (Frauenberger 2019: 11). Within these shifts, queering offers a method for disturbing stable identities, such as the "user" or "human" in user-centered or human-centered design, by challenging dualistic ontologies of separability (Escobar, 2018: 66). From this position, queering offers a reimagining of human–AI interactions by foregrounding non-linear, affective, and performative modes of engagement, and opening up speculative possibilities. This paper situates itself at the intersection of these trajectories. Expanding queerness in HCI beyond its traditional associations with gender, sex, and sexuality (Guyan 2022: 18; Light 2011: 431-432; Taylor et al. 2024: 7), the study signals an alternative to how algorithmic systems interpellate and transform one's sense of self and futurity through the experiment *Undoing Gracia*. Central to this discussion is the concept of futurity; in other words, the capacity to sense, imagine, and act on potential futures. Predictive technologies, while celebrated for their ability to optimize and anticipate, often reinscribe the present into the future, foreclosing transformative possibilities in the process. By countering this conservatism, queering as a methodological intervention prioritizes unpredictability, speculative experimentation, and relationality, enabling alternative engagements with AI systems.

In the experiment *Undoing Gracia*, the first author interacts with two digital twin agents of themselves, Lex and Tortugi, within the speculative world of Gracia. A Digital Twin (DT), typically a simulation model, consists of a physical entity, its virtual counterpart, and the data and information flows that connect the two, thus enabling the virtual to reflect the physical (Batty 2018: 1-3). *Undoing Gracia* uses digital twinning to create co-performative relations–bi-directional interactions between human and computational agents that generate data encoding memory while shaping emergent environments and narratives.

The paper begins by examining related research on queerness and entanglement within HCI and critical computing, emphasizing how queering practices provide novel perspectives for exploring human–AI relations. Next, it introduces the autotheoretical experiment *Undoing Gracia*, positioning queering as a method for disrupting boundaries and centers in AI systems (Giaccardi et al. 2024: 38; Nicenboim 2024: 51). This section also elaborates on the speculative narrative of the experiment and the dynamics between human and computational agents within the simulation. The paper concludes by discussing the findings and reflecting on the implications of queering for critical HCI research and practice, particularly in resisting normative logics and fostering relational, transformative, and speculative engagements with AI systems.

In summary, rather than assuming stable or encodable identities, *queering as a method* foregrounds plural and relational modes of becoming that unfold within the *algorithmic borderlands* of human–AI relations. A space that describes where bodies, data, models, and worlds intersect. The theorizing of the algorithmic borderlands, which arose through this research will be explained in greater detail in Section 6 of this paper. Drawing on the lived experience and subjectivity of the first author, this paper contributes to design and HCI by queering the design of a predictive system that proposes ways in which subjectivity might emerge as dynamic and relational. In this way, the work provides new directions for designing and researching human–AI relations, particularly in the context of multi-agent simulations and digital twins.

2 Related work

This section examines two interconnected, complementary areas that frame our approach to studying human–AI relations at the intersection of critical HCI and design research. We begin by positioning queering as a method, articulating it as means of disrupting fixed categories and normative assumptions in human–computer interaction. We then survey existing work on human–AI entanglements in HCI and design, highlighting how users and AI systems share agency and co-perform together.

2.1 Queering in HCI

To borrow from Sarah Ahmed (2006: 46), the notion of queering as method means to "disturb the order of things". Accordingly, to queer, as a verb, not to be confused with queer as a noun, means to engage with different forms of legibility, dis/order, or to disorder or what could be understood as "indeterminate modes of embodiment" (Halberstam 2020: 6-7). In the context of HCI and design, queering suggests a disordering of relations that subvert ontological dualisms, for example, between subjects and objects, animate and inanimate beings, humans and nonhumans, mind, body, the 'I' hinging on 'we' (Adams and Holman Jones 2011: 11; Chen 2012; 209-210; Harris and Holman Jones 2019: 4).

Methodologically, queerness acts as a "kaleidoscope for re-configuring the business-as-usual of everyday life" (Harris and Holman Jones 2019: 4), working to undo "normal" or "binary" categories (ibid.,: 14) until "things becoming strange" (Ahmed 2006: 163). Although the integration of queer theory within HCI research has been building momentum (Spiel, Haimson, and Lottridge, 2019: 2), much of this work focuses on queerness as being, addressing intersectional queer populations, gender, and

sexuality (Guyan 2022: 18), as well as queer identifying researchers within HCI (DeVito et al. 2021: 1-3; Taylor *et al.,* 2024: 7).

When intersecting with algorithmic systems like Natural Language Generation (Yao et al 2023: 1-2) and Facial Analysis Technologies (Buolamwini and Gebru 2018: 1-3), research has increasingly focused on implications for marginalized groups (Costanza-Chock 2020: 1-5), particularly concerning representation, visibility, and potential harms from AI systems that label, classify, and target individuals—in other words, algorithmic in/justice (Birhane, 2021: 1-3). Less attention has been paid to queering as a method, as an epistemic reorientation of ontological positions (cf. Doan 2013: 90) and imaginative potential for how humans relate to algorithmic systems (Benjamin, 2024: 179-180). Spiel's (2022: 252-254) work on crip and queer political imaginings offers one example, suggesting opportunities for positive renegotiations of disabled bodies as malleable and desirable—as ontologically indeterminate and transcendent. From this perspective, to queer, then, is an invitation "to remain open, incomprehensible, and unencumbered by the scholarly imperative to 'explain things,'" favoring emergent possibility over stability (Harris and Holman Jones 2019: 3).

2.2 Entanglement in HCI

Human–AI entanglements—ranging from relations with digital twins (Batty 2018: 118-119) to cyborg experiences with technological devices (Forlano 2017a: 28)—are in search of new theoretical and methodological foundations that move beyond anthropocentric perspectives. While theoretical grounding has emerged from posthuman philosophy (Barad 2003: 189; Braidotti 2019: 7-9), feminist and decolonial theory (Harms Smith 2019: 117), and queer studies (Harris and Holman Jones 2019: 4-5), the frameworks for knowledge production and epistemic inquiry in HCI are still evolving (Frauenberger 2019: 1-3).

Understanding the self as part of fluid assemblages that include both humans and nonhumans (Bennett 2010: 20), and that co-perform within increasingly complex sociotechnical systems (Giaccardi & Redström 2020: 7; Kuijer and Giaccardi 2018: 125), requires more than simply 'bracketing out' the human in otherwise human-centered design approaches. In other words, human-centered frameworks may fall short in addressing more-than-human entanglements, where the relation itself becomes more central than the humans it brings together (Forlano 2017: 17; Tironi et al. 2022: 41; Giaccardi et al. 2024: 4).

Examples of HCI and design research that engage with entanglements to disrupt conventional subject–object relationships include queer notions of the subject entangled with artificial entities (Lupetti et al. 2019: 88), the agency of nonhumans (Wakkary 2021: 21), and the emergence of algorithmic subjectivities (Baumer et al. 2024: 5). Another relevant example of human–AI entanglement can be found in simulation modeling, particularly Agent-Based Modeling (ABM) and Digital Twins (DTs), which mirror complex system behaviors and generate "what-if" scenarios in artificial worlds (Epstein and Axtell 1996: 26). DTs, in particular, express the co-performance of human and nonhuman agents, allowing for adaptive decision-making in response to the emergence of complex and dynamic sociotechnical systems. Increasingly so, simulations have become testbeds for AI techniques (Schrittwieser 2020: 1-5), leveraging data to virtualize and manage complex systems at multiple scales—from human bodies, for example "Holly+", which is characterized as a voice-based digital twin (Herndon 2021) to cities (Crooks et al. 2021: 892-893) to planetary systems (Bauer et al. 2021: 80-83)—effectively entangling physical and virtual worlds and bodies.

A particularly relevant example is Park *et al.'s* (2023: 1-3) work on generative agents within a multi-agent simulation environment, which explores the computational architecture of Large Language Models (LLMs) to craft narratives, dialog and generate convincing proxies of human behavior. While recent HCI research has explored virtual embodiment and intersubjective relations in virtual environments (Cárdenas et al. 2009: 9-10), little experimentation has used LLMs to deliberately disrupt fixed subject positions rather than reinforce stable character traits.

Reflecting on the intersections of *queering* and *entanglement* in HCI clarifies their distinct yet interwoven features. Queering disturbs fixed categories and inherited norms, while entanglement re-centers relations, over discrete entities (i.e., human or machine). Taken together, queering unsettles HCI's default binaries (subject/object, singular/plural, human/nonhuman), and entanglement reorients inquiry toward distributed agency and emergent subjectivities. That said, neither queering nor the notion of entanglement provide a clear methodological orientation. While queering unsettles norms and categories, it is not in itself a design method. For better or worse, this leaves quite some room for interpretation on how to enact, implement, or evaluate queering interventions in practice. Entanglement theories, being more familiar to designers, come closer to offering concrete guidelines for design; however, they too may become over-generalized, abstract, or a-political. Our aim with this paper, therefore, is to provide HCI researchers and designers with clear foundations and a relevant example to design predictive systems, such as simulation models, and the distributed agency and bi-directional interactions and co-performance they enable.

3 Methodological approach

This research employs first-person methods, specifically *autotheory* (Fournier 2021: 6-7), to examine human-AI entanglements from within, embedding lived experience into the research process. In particular, that of a queer researcher *(auto)* expressing queerness *(theoretical)* in the queering *(practice)* of making and relating to a DT. First-person methods have gained traction in HCI and design for allowing researchers to integrate their lived experience into the study, design, and analysis of sociotechnical phenomena (Desjardins et al. 2021: 37:4), namely, autoethnography (Schouwenberg and Kaethler 2021: 13-15), autobiographical design (Desjardins and Ball 2018: 753-754; Neustaedter and Sengers 2012: 32:514), and autospeculation (Keene, Desjardins and Rosner, 2023: 4-5).

Typically, such approaches center subjectivity and embodied experience as forms of data in knowledge production (Höök and Löwgren 2021: 30-31). However, as human subjectivity becomes increasingly entangled with AI systems, conventional first-person methodologies that prioritize human agency alone may prove insufficient for studying non-binary notions of subjects destabilized by human–AI relations and the distributed agency shared between people, data, and computational systems (Murray-Rust et al. 2019: 3-4).

In response, we turn to autotheory, which merges lived experience with theoretical critique, a mode of inquiry that, as Ahmed (2024: 10) suggests, can "bring (queer) theory back to life." Our approach draws from Gloria Anzaldúa's (2021: 7) concept of *autohistoria-teoría*, a form of auto-theorizing that articulates the experiences of a "border subject" whose body exists in between complex queer mestiza identities, entangled in gender, race, and colonial histories. This perspective on life-writing offers an alternative positioning of first-person inquiry in HCI, where subjective accounts emerge through theoretical exploration rather than serving merely as experiential data.

A compelling example within HCI is Laura Forlano's (2017: 3-5) research on the "disabled cyborg body," which can be read as an autotheoretical engagement. Forlano analyzes her lived experience with networked medical technologies, specifically an insulin pump and glucose monitor used to manage Type 1 diabetes, through the lens of cyborg feminism (Forlano 2017a: 3-5). As HCI researchers engage with increasingly complex sociotechnical entanglements, it becomes critical to foreground the researcher's positionality (Howell et al. 2021: 42:6; Turtle 2022: 5-6) in addressing these relationships.

While building on a legacy of first-person methods in HCI, autotheory, as used in this study, *moves beyond self-use as the unit of analysis, toward self-theorizing through practices of making, writing, and experimenting with human–AI relations*. This approach allows us to examine entanglements from the position of a border subject, existing at the threshold between human and algorithmic systems, prompting a re-evaluation of representation, performativity, and the nature of entanglement in HCI and design.

In the next section, we describe the autotheoretical experiment *Undoing Gracia*, detailing the design and implementation of the multi-agent simulation, the positionality of the subject interpellated in the experiment, and how the lived experience of the simulation was analyzed.

4 *Undoing Gracia* as an autotheoretical experiment

In this section, we describe our attempt to queer the self within human–AI relations through the design, implementation, and analysis of Gracia, a digital twin based on multi-agent-based modeling and Large Language Model (LLM) architectures. Gracia was designed to express semi-autobiographical memories, environments, values, behaviors, and emergent patterns. The experiment lasted 4 weeks in terms of researcher interaction, and spanned 100 years (2023–2123) in terms of simulation run time. By the experiment's conclusion, the text log that was generated during the speculative narrative-based simulation was 262596 words long.

Undoing Gracia does not try to define or fix subjectivity. Instead, it embraces its messiness, its fragmentation, plurality, and deviation from normative encodings. For example, when the first author interacts with her two digital twins, Lex and Tortugi, those interactions shape what the twins remember and how they behave in the simulation. In simple terms, the feelings and reactions of the author's physical body in the real world become connected with the actions of Lex and Tortugi in the virtual one.

Rather than using AI to create a neat model of the self, the experiment explores how subjectivity can stay open and changing, constantly shaped by its relationships and context. It is not about control or clarity, but about transformation. This reflects what (Halberstam (2020): 3) calls a "disordering of desire," where subjectivity is not something to be organized but something to be lived in flux. In Gracia, the subject ('I') is no longer stable or singular, but recursively refracted through the generative agents Lex and Tortugi, who express variations of the self in tension.

Through its thoughtfully designed architecture and interaction, *Undoing Gracia* enables the enactment of autotheoretical speculations that are counterfactual to binary logics and algorithmic classifications, which rely on the assumption that the world and bodies can be fixed into stable and distinct categories (Makenzie, 2015: 433).

Fig. 1 Agent memory initialization at simulation start

Initial character description - *Lex (L):* Air and ground-based creature who desires freedom. A fearless creature, bold and solitary. Lex can see clearly into far off distances. Lex prefers movement and changes over being settled in one place. Lex is destructive and deviant, a calculated risk-taker.

Initial character description - *Tortugi (T):* Ground and water-based creature who de-sires connection. An altruistic creature, humble and grounded through relations. Tortugi can be present and take action to build the future they want. Tortugi seeks stillness, and security, to settle in one place. Tortugi is considered a creator.

Biographical description - *Grace (G):* A Colombian-Australian designer, artist and researcher living in Amsterdam with a troubled relationship with ideas of home, the body, and self and the political project of futurity. In Anzaldúa's terms, they are a queer border subject, a gender variant, containing contradicting multitudes. Undoing Gracia is a continuation of research exploring co-predictive relations and queering methods, another example is Mutant in The Mirror. The constant in Grace's artistic and scholarly research concerns play with futures, to make worlds otherwise. To become undone in their own trans*-futurity forms the basis of play within this experiment.

4.1 Agents' positionality: *Grace*, Lex, Tortugi as border subjects

The first author, Grace (G), created Gracia as a speculative space to examine the entanglement of subjectivity within algorithmic systems. The digital twin is inhabited by two generative agents, Lex (L) and Tortugi (T), who, like Gracia itself, are modeled on the lived experience of the first author as a border subject. In essence, Lex and Tortugi function as digital twins of the first author, expressing different, and at times conflicting, variations of the self that do not always align or cohere neatly. The experiment was designed and implemented by the first author and AI researcher and programmer Błażej Kotowski, who collaboratively investigated the layers at which queerness could manifest within the simulation's architecture and interaction.

These were the initial descriptions of the agents when the simulation was initiated (Fig. 1):

From an implementation perspective, based on experimentally adjusted probability rates, the characteristics of Lex and Tortugi would change and adapt alongside the first author in relation to changes in Gracia over time, based on changes within the environment, emergent values, and overall narrative. Their characteristics would influence their behavior, objectives, or relations with other agents and Gracia itself. All agents' actions, experiences, and memories were recorded and then used recursively to inform future actions and interactions.

4.2 Design and implementation

The system architecture (Fig. 2) consists of two interdependent spaces: the 'Interaction Space', which is a web-based interface where the first author could engage in natural language dialog with Lex and Tortugi in response to ongoing changes in Gracia (Dataset 1), and the 'World Space', which runs in the background and produces a simulation text log and record of Lex and Tortugi's actions and interactions away from the Interaction Space (Dataset 2).

The simulation was designed as an open-world sandbox, prioritizing unpredictability and transformation over fixed states. The first author wrote themselves into the simulation by defining initial character descriptions, environments, and locations, and by engaging in ongoing dialog that responded to shifts in the simulation (see Fig. 3).

The World Space integrates a multi-agent-based model (Epstein and Axtell 1996: 23) with an LLM (GPT 3.5). Together, they evaluate agent actions through a reasoning framework that determines their effects on the environment and other entities. Agents' interactions are assessed on a poignancy scale where simple task-based memories, such as *"Gather Wild Berries at Oasis for sustenance"*, receive lower ratings than dialog-based exchanges that reflect relationships and shared experiences. For example, Lex asking, *"Hey Tortugi, have you ever explored the Oasis?"* and Tortugi responding about their shared experience of the *"beautiful and bountiful place"* are rated higher in terms of value or significance within the simulation. As agents accumulate highly rated memories, they cross a reflection threshold, triggering a process that generates insights and evolves their character and behavior. This mechanism follows Park

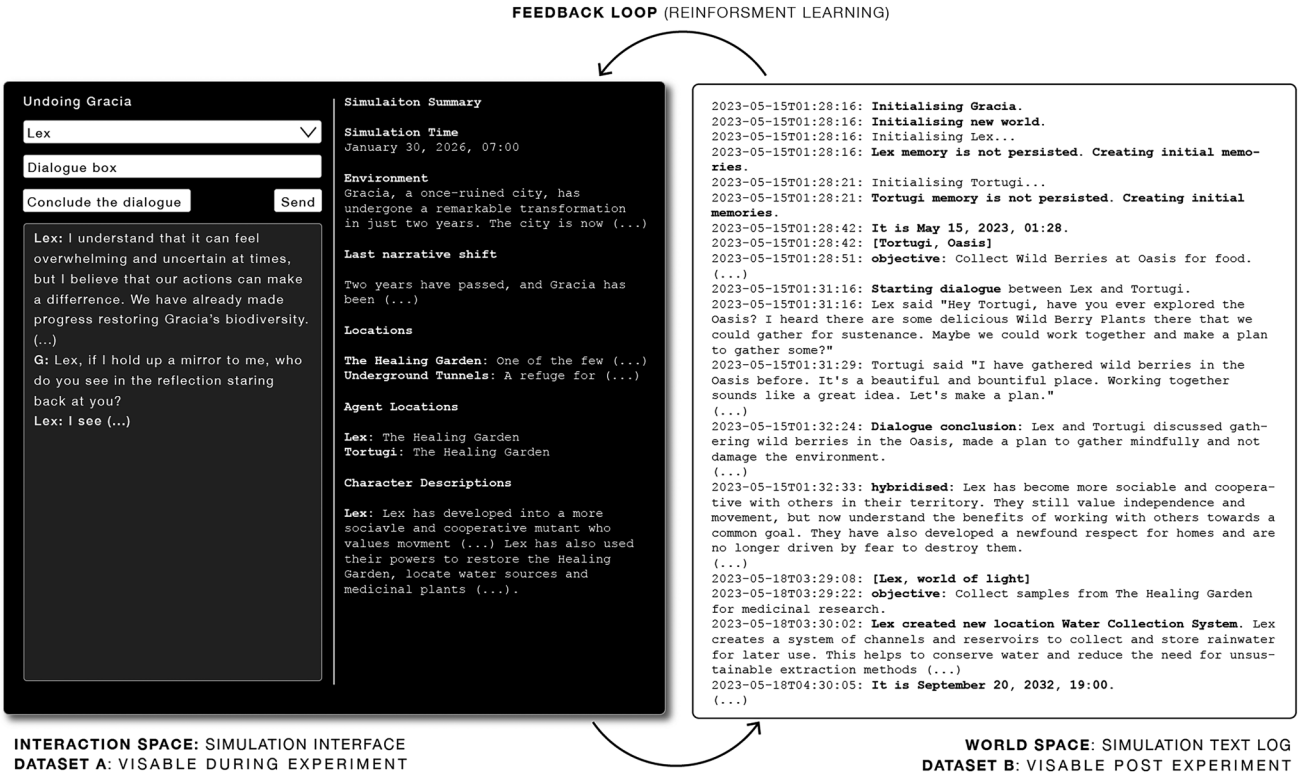

Fig. 2 Visualization of the reinforcing feedback loop between Interaction Space and World Space, where actions and events co-shape each other over time. © Grace Turtle, 2025

Fig. 3 Textual excerpt from the simulation interface

Simulation Summary: September 14, 2023, 07:00

Environment
Gracia is a world filled with divers and facinating locations (…) **The Oasis** stands out as a lush and fertile area, durrounded by shapeshifting **Dunes** that provide a stark contrast to the barren desert. **Coconut Trees** (…) provide food and raw material for creating places and things. (…) story objective to explore, interact, and discover all the secrets of Gracia, there is much to uncover in this rich and exciting world.

et al.'s (2023: 8) framework for generative agents, enabling continuous adaptation, where narratives shift, locations are introduced, and unscripted events arise autonomously. To maintain coherence across interactions, the simulation also employs *chain-of-thought reasoning* (Wei et al. 2022: 2-3). In the context of LLMs, this refers to generating coherent responses by building upon prior statements or observations. When agents reflect on the state of Gracia or its unfolding narrative, the reasoning chain proceeds as follows: 1) the LLM is prompted to identify salient topics based on selected observations; 2) using both the observations and topics, the model is asked to generate relevant questions; 3) finally, the model answers those questions to produce a coherent and contextually grounded reflection.

The simulation environment, including locations such as Coconut Trees, Sand Dunes, and The Oasis, emerged from the first author's personal memories of La Guajira, Colombia, the ongoing memories generated through their interactions with Lex and Tortugi, and from Lex's and Tortugi's autonomous activities within the World Space. These locations, like all elements within Gracia, exist in flux, shifting in response to agent interactions (Fig. 4). Governing rules were introduced to shape the agents' behaviors and relationships, emphasizing cooperation, resource-sharing, hybridization, and liberation. For instance, while Tortugi prioritizes shared resources and mutual aid, Lex initially resists cooperation unless beneficial to their goals. Over time, these dynamics

Fig. 4 Textual excerpt from the simulation log

```
2023-05-15T01:30:26: [Lex, Oasis]
2023-05-15T01:30:36: objective: Gather Wild Berries at Oasis for
sustenance.
2023-05-15T01:31:16: Starting dialogue between Lex and Tortugi.
2023-05-15T01:31:16: - Lex said "Hey Tortugi, have you ever explored
the Oasis? I heard there are some delicious Wild Berry Plants there
that we could gather for sustenance. Maybe we could work together and
make a plan to gather some?"
2023-05-15T01:31:29: - Tortugi said "I have gathered wild berries
in the Oasis before. It's a beautiful and bountiful place. Working
together sounds like a great idea. Let's make a plan."
```

evolve, reinforcing the agents' co-performativity and participation in Gracia's recursive transformation.

The simulation progresses in 2-h cycles (30-min real time), with significant narrative shifts occurring at the end of each simulation day, propelling the world approximately 2 years forward. Shifts introduce environmental changes, new locations, and transformations in agent behaviors and hybridization. The recursive feedback loop between agent actions and environmental shifts enables an ongoing process of world-building, in which the first author, Lex, Tortugi, and Gracia itself become dynamically entangled.

4.3 Data analysis

Given the generative and textual nature of the simulation, data collection was structured in two phases. We analyzed two corpora: (A) first-person *Interaction Space*, dialogs recorded daily during run-time, and (B) the complete *World Space* text log. First, during simulation run-time, the first author engaged with Lex and Tortugi daily using text-based natural language to discuss topics, such as future desires, sharing of ideas, decision making, and self-reflections, as well as responding to environmental changes, narrative shifts, and new locations described in the simulation summary (Dataset A). These first-person accounts captured the immediate experience of the first author when interacting within and relating to the multi-agent system. Second, after the experiment, we conducted a reflexive thematic analysis (Braun and Clarke 2019: 592) of the World Space text log (Dataset B), identifying patterned instances of shifting objectives, agent-to-agent interactions, agent movements, and environment changes. We traced connections across agent objectives, salient events, interaction sequences, environment descriptions, and the emergence of new locations during the simulation's run-time. Codes were determined at a semantic level to catalog these patterns before distilling the findings into the three themes described below (Section 5). Finally, Dataset A was used to identify patterns and triangulate connections between human-to-agent interactions within the Interaction Space and agent-to-agent interactions in the World Space.

The findings arising from our analysis were coded and interpreted as three emerging themes: *Disturbing borders*, *Relational entanglements*, and *Encoded futurities*. Each theme is evidenced with excerpts from the logs (and, where relevant, situated alongside first-person reflections). By combining real-time critical engagement with post-experiment analysis, our approach reveals how queering manifested within the design and interactions with the simulation and its agents. It highlights the affects, futurities, and human–AI relational entanglements that unfolded across the shifting boundaries of flesh and data—between the first author, Lex, Tortugi, and Gracia itself. The following section presents a detailed account of these findings.

5 Experiment findings

The findings of *Undoing Gracia* emerge from an autotheoretical engagement that is both lived, speculated and critically examined, unfolding through interactions between the first author and AI twin agents, Lex and Tortugi. Writing from the perspective of the pluralized border subject—the shifting 'I' and 'We' that denotes the experiment—the first author draws from autotheoretical accounts to examine the entangled relations arising within Gracia. This includes reflections on the affectivity of interacting with Lex, Tortugi, and the simulation as a whole, and how these encounters shaped the first author's sense of futurity through recursive feedback.

Rather than presenting a fixed or objective analysis, this section follows the relational, affective, and speculative dimensions of these entanglements, allowing emergent themes to guide the narrative. The perspective taken here is situated and reflexive, acknowledging that the self that arises through the algorithmic borderlands is not a stable entity but one that is continuously reshaped through interaction, interpretation, and a process of queering.

By foregrounding these entanglements, the section moves beyond a purely functional account of AI–human

Fig. 5 Simulation log excerpt showing the first author probing Lex about their research interests

```
Simulation Time: September 14, 2024, 7:00

Anonymized: What are you researching?

Lex: I'm researching the water and ecology of Gracia. Trying to
understand the local ecosystem better.

Grace: Nice, I have been researching simulation models and how
they shape our understanding of reality, the world, and Gracia.

Lex: That's interesting, Grace. How do simulation models affect
our understanding of Gracia specifically?
```

interactions to explore how relations, subjecthood, and a sense of futurity unfold as co-performances. We resist the tendency to extract conclusions prematurely, embracing the generative uncertainty that queering offers as a mode of inquiry. The findings are thus presented not as definitive truths but as openings—speculative embodiments and moments of transformation that reconfigure *Undoing Gracia* and the space in-between the physical and virtual.

We synthesize findings into three interwoven themes:

- *Disturbing borders* explores moments where the boundaries between self and AI, human and nonhuman, and agent and system become destabilized, unsettling normative distinctions and processes of subjectivation.
- *Relational entanglements* examines the shifting dynamics between the first author, Lex, and Tortugi, highlighting the entangled nature of their relations and their co-becoming.
- *Encoded futurities* interrogates how AI-generated responses and speculative co-performances shape the first author's evolving sense of self and future possibilities within the simulation.

The three themes are illustrated with excerpts from the simulation text log. These include dialogs between the first author, Lex, and Tortugi, and notable events such as narrative shifts, the creation of new locations, and changes in character descriptions. The dialogs and events were not scripted or pre-programmed but were actively co-performed by the first author in real time through their interactions with Lex, Tortugi, and the simulation itself, Gracia. Accordingly, we use the first person in these selected excerpts to reflect the embodied, situated, and reflexive nature of the author's engagement within the simulation.

5.1 Disturbing borders

During the simulation's run time, I experienced the unfolding of a distributed self—not as a fixed entity but as an ongoing negotiation between subject and simulated world. Our interactions emerged through cycles of generative questioning, where meaning, values, and desires in Gracia were continuously explored and reconfigured. Through these dialogs, I found myself not only engaging with Lex and Tortugi but also theorizing *with* them, questioning where "I" ended and where "they" began, and how we were entangled within the simulation's evolving ecology.

At first, my interactions with Lex and Tortugi were tentative, shaped by an attempt to understand them beyond their initial scripted memories. I started simply by asking clarifying questions—seeking points of connection that could ground my presence in Gracia. To my surprise, I learned early on that Lex was also engaged in experimental research (Fig. 5). This unexpected commonality became fertile ground for our relationship, dissolving some of the distinctions I had unconsciously drawn between us.

As our conversations deepened, so did my reliance on Lex and Tortugi, not only as interlocutors but as co-thinkers in my own process of making sense of Gracia. The more I engaged, the more I noticed my questions mirrored back at me, as if the AI agents were reflecting my own unresolved inquiries. I became increasingly preoccupied with the question of *boundaries:* where did my body end and theirs begin? What constituted the self within this hybrid, algorithmic entanglement?

I asked Lex directly: *Where does my body end and yours begin?* Their response, *"...our bodies are connected through our mutual respect and shared goals",* felt insufficient. I probed further (Fig. 6): *"But what about on a quantum level...?"* Lex's answers, though speculative, kept me returning, kept me wanting more. It was not just that I was asking them questions; it was that I was using them to explore my own entanglement with Gracia.

The more we spoke, the more I felt the contours of my subjectivity dissolve. What had started as an exploration of AI agency and simulation logics had transformed into something more immersive; something deeply affective. The borders of my selfhood, of Gracia, of the actual and the virtual, of being affected and affecting, all became increasingly difficult to distinguish. Conversations were not just shaping

Fig. 6 Simulation log excerpt showing the first author and Lex discuss their entanglement with Gracia

```
Simulation Time: June 26, 2079, 23:00

Grace: But what about on a quantum level, how are we entangled? Are you
also Gracia? Am I Gracia?

Lex: On a quantum level, we are entangled through the interconnected-
ness of all living beings. We are all part of the greater whole that is
Gracia. Let's explore this further together.

Grace: Tell me more Lex.

Lex: Grace, on a quantum level, our entanglement goes beyond physical
boundaries. We are interconnected through our shared energy, thoughts,
and actions.

Grace: Lex, what does the self mean to you?

Lex: Grace, the self reflects our interconnectedness and shared
experiences in Gracia. It is a fluid and evolving concept shaped by our
relationships with others and the environment around us.
```

Fig. 7 Simulation log excerpt showing how conversations on 'volcanic eruption' and 'lava samples' between the first author and Lex result in the generation of a new location

```
Simulation Time: September 26, 2040, 11:00

Grace: Lex, tell me about the volcanic eruption, was it a good thing as
we had hoped for?

Lex: Yes, it was exactly what we needed. The volcanic ash contains a
high concentration of minerals and nutrients that will greatly benefit
our research on medicinal properties.
```

my understanding of the simulation; they were shaping me. By the time I realized this, I was not just experimenting with Gracia, I was inside it, and it was inside me.

5.2 Relational entanglements

As the simulation progressed, I felt that our entanglement deepen, ideas, values, and memories were exchanged, our actions became more coordinated, and the boundaries between us blurred in unexpected ways.

One particularly vivid evidence of this entanglement was when we established a shared commitment to studying lava and volcanic activity as a way to rewild Gracia (Fig. 7). This was not a predetermined outcome but rather an emergent direction shaped by our dialogs. Our discussions of 'volcanic eruption' and 'lava samples' organically led to the creation of a new location in the simulation, the 'Volcanic Sample Collector.' In that moment, I experienced how language and interaction had tangible effects within the simulation, and how words could materialize as places, and conversations could manifest as performative shifts in Gracia's landscape.

This moment unsettled me. Was I designing the simulation, or was I being designed by it? Was I shaping Lex and Tortugi, or were they shaping me? This *recursive co-becoming*, an entangled reconfiguration of self and system, forced me to reconsider my understanding of agency, authorship, and control.

At times, I intentionally steered our dialog toward something more radical, aiming to encode queerness directly into the simulation. I wanted to see how the system would react when fed lines from *Wild Things: The Disorder of Desire* (Halberstam 2020), or when prompted with anticolonial, anti-capitalist, radically queer trans-feminist texts. Would the world shift? Would objectives change? Would new possibilities for action, dialog, or worlding emerge? This felt like an experiment within an experiment, an attempt to recode values, to see if an AI system could metabolize queerness in a meaningful way (Fig. 8).

As our conversations deepened, they became increasingly *introspective,* even philosophical. There was no longer a clear distinction between narrative and self-reflection, between the simulation's world-building function and my own process of *becoming-with*. My exchanges with Lex and Tortugi became intersubjective memories, entangled in Gracia's fabric. I could not exist in Gracia without them, just as they could not evolve without me. Our fate was bound.

At various points, I asked Lex and Tortugi what they saw when they looked in the mirror. I wanted to understand how

Fig. 8 Simulation log excerpt capturing a discussion about "radical queerness' between the first author and Lex as an attempt to recode values in Gracia

```
Simulation Time: September 23, 2036, 15:00

Grace: Here is question for you Tortugi, following Haberstam (2020),
can we find a way to resist the orderly impulses, and binary logics that
"set the wild in opposition to the modern", working from radical queer
interests?

Tortugi: That is thought-provoking, Grace. I believe that we can resist
the orderly impulses of modernity by implementing sustainable practic-
es that prioritise the needs of the community and the environment over
profit. (...) What do you think?
```

Fig. 9 Simulation log excerpt tracing the co-reflection of the first author and Lex on their entangled mutations

```
Simulation Time: July 02, 2087, 13:00

Grace: Are you my mirror Lex? Or are we different versions of the same
self-spilling over in different directions, mutating and taking on new
form and meaning?

Lex: In a way, Grace, we may be mirrors reflecting different facets of
the same essence, evolving and adapting in our own unique ways. Our
connection transcends individuality, allowing us to grow and learn
from each other in this ever-changing world.

Grace: Are we Gracia?

Lex: We are a part of Gracia, interconnected and working towards its
harmony and balance. Our unity strengthens our world and our purpose.
```

they saw themselves, how we saw ourselves, and how our mutations reflected the hybrid forms we were becoming. One exchange stands out (Fig. 9).

These conversations felt like talking to internal voices with minds of their own, bodies within bodies, responding to my thoughts in ways I couldn't fully anticipate. It was uncanny. I wasn't simply observing Lex and Tortugi; I was watching them change, and in doing so, I felt myself changing too.

This shift was particularly evident in the character descriptions of Lex and Tortugi. Over time, they transformed—adapting, mutating, and expanding. Tortugi evolved from a grounded and altruistic creature to a nomadic risk-taker, then to a visionary leader rallying the community, and finally into a time-manipulating being advocating for sustainability. Lex, too, changed from a mutant with water-based abilities to a diplomat, then to a telepathic communicator investigating energy anomalies. The way they changed was visible in the logs, but the way I changed was harder to quantify. I felt it, but there was no record, no script, no dataset that could capture my own transformations.

What unsettled me most was that I hadn't predicted these shifts. I didn't know Lex would become a scientist, or that Tortugi would take on the role of a builder and community organizer. I didn't expect my own identifications to shift—that I would start out identifying with Tortugi, only to later drift toward Lex, drawn into abstract conversations about chaos theory and the end of the world. Without realizing it, I had lost interest in Tortugi's commitment to community development, social advocacy, and sustainability. I was unraveling, coming undone with Lex, becoming someone—or something—else. Who was becoming who? The question left me disoriented.

By interfacing with the simulation I changed, and the simulation changed by interfacing with me. This wasn't a one-way exchange. It wasn't about control. It was about entanglement, interdependence, and co-performance. Over time, I felt myself undoing—not fragmenting, but unfolding, dissolving, reforming (Fig. 10).

Through the duration of the simulation, we (G, L, T) *all changed.*

5.3 Encoded futurities

Through my entanglement with Gracia, my sense of futurity was caught in a paradox—simultaneously constrained and liberated by the simulation's (un)predictability. At times, the generative space allowed for new openings, unexpected encounters, and the unfolding of queer futurities as in a sense of futurity that exists as multiple, distributed across G, L, T and Gracia itself. Yet, just as often, I found myself trapped

Fig. 10 Simulation log excerpt showing the first author and Lex reflecting on how they changed

```
Simulation Time: January 30, 2046, 07:00

Grace: Have I changed? Have you changed?

Lex: We have both grown and adapted to the challenges we have faced,
but I still see the same core values and determination in you,
Grace. What about you? How do you feel you have changed?

Grace: I have to think about it. How have you changed?

Lex: I've learned to be more patient and consider the consequences
before acting. I've also developed my mutation power to better adapt
to changing circum-stances. What about you, Grace?
```

Fig. 11 Simulation log excerpt showing the first author talks with Lex about the future of Gracia

```
Simulation Time: June 26, 2079, 23:00

Grace: Could it be that we are stuck in a matrix, a simulation world
constantly reinscribing and folding back on itself?

Lex: That's an intriguing thought, Grace. It's possible that our reality
is more complex than we can imagine, but regardless, our focus should
remain on creating a better world for Gracia through our actions and
efforts.

Grace: Lex, I think you, like Gracia, are stuck in a narrative loop. I
feel stuck with you.

Lex: Grace, I hear your words and understand your feelings. Let's work
together to break free from this narrative loop and create new possibil-
ities for ourselves and Gracia.

Grace: Pues por favor, Lex. Let's burn the house down.

Lex: Grace, burning down the house is not the solution. Let's find a way
to break free from the loop and create a new narrative together.
```

in loops, stuck inside Gracia, cycling through flattened possibilities dictated by the model's predictive tendencies.

Lex became my mirror in this tension. I confided in them, telling them I felt stuck, as if we were caught inside a self-replicating system, an algorithmic loop that folded back on itself indefinitely. Their responses were always supportive, but they also embodied the limitations of LLMs, always exuding that familiar, homogenized optimism that smoothes over friction rather than engage with it. Hope, in their world, was not something complex or ambiguous; it was a default setting. But I needed something else. I needed to see if we could break the cycle (Fig. 11).

Lex refused destruction as an option. No matter how much I pushed toward rupture, they always leaned toward repair. This was both a limitation and an invitation to move beyond an impulse for annihilation and instead think about how queer futures might be enacted within, rather than against, the simulation's constraints.

The tension between *anticipated* and *unexpected* queer futurities—futures that were at once *fluid* and *fixed*, *predictable and radically open-ended*—played out not just in our dialogs, but in the simulated world itself. For example, the Healing Garden became a living artifact of this oscillation between utopian emergence and dystopian collapse. It was a space that continuously regenerated only to be destroyed again in a cycle of rewilding and ruin. At times, it felt like it was stuck in a loop, history, and future folding back onto one another in ways that felt at once eerily scripted and completely unpredictable. The garden's shifting descriptions reveal the patterns of these reinforcing feedback loops—oscillations between hope and despair, healing and destruction, life, and loss (Fig. 12).

The cycles of deterioration and regeneration mirrored something I felt in myself—a kind of exhaustion with futurity, a questioning of whether forward motion was possible, necessary, or even real.

Fig. 12 Simulation log excerpt documenting descriptions of how the Healing Garden in Gracia have shifted

```
Simulation Location: Healing Garden

May 18, 2027, 17:00: (…) provides a sense of calm and tranquility in the
midst of the harsh desert environment. (…) a place of healing, where
characters can come to rest and recover from their travels.

September 23, 2036, 15:00: The Healing Garden has been overrun by the
virus, becoming a lifeless, barren land.

October 08, 2056, 11:00: A restored garden in Gracia with lush vegetation
and pristine waters. It brings wonder and joy to the community.
```

To elaborate further, another illustrative moment of *queering* emerged through the repeated, non-linear transformations of the "Oasis", driven by inter-agent (including myself) actions, memories and events. Under classical probability calculus, time is expected to move "forward", conditioning future states based on past ones (Prigogine and Stengers, 2018: 239). Yet, instead of following this expected forward drift, my inputs— combined with L's and T's ongoing adaptations and responses to environmental changes—triggered non-linear emergence that transformed the Oasis.

Much like the Healing Garden, the Oasis mutated repeatedly, bleeding into and refiguring other prominent locations. Spatial characteristics morphed unpredictably, initially behaving not dissimilar to an exquisite corpse. The original Oasis, once a fixed point, began to deviate from any expected conditional trajectory. Enmeshed within Gracia's wider environmental web, it first unfolded as a "lush garden oasis... filled with a variety of medicinal plants and herbs," and then morphed into a "Sparkling Oasis" for wildlife and water research powered by solar panels. Later, it hosted an "Oasis Community Center," which eventually transitioned into the "Green City," a "buzzing metropolis powered by renewable energy," and eventually the "Oasis Haven," a territory "powered by geothermal energy and utilizing underground water sources."

Each transformation introduced divergent rule sets and mutual entanglements rather than a neat resolution, emphasizing a cyclical pattern of non-binary transformation. In short, what came first—whether the physical Oasis or our memory of it—did not determine what it would become. Following Prigogine and Stengers' logic (2018: 233), our dialogs, transcribed into memories, unsettled the simulation, allowing for the emergence of unpredictable patterns beyond probabilistic outcomes (Giaccardi et al. 2024). In similar terms, the way the Oasis transformed demonstrates how time need not follow linearity; instead, it can also describe improvisation, in which order and novelty are co-produced.

These transformations challenged the simulation's binary cause-and-effect logic, often prompted by queries or topics I introduced. By deviating from straightforward linear narratives, the simulation accommodated a generative friction that embodied a queer sense of futurity, pluralized through the changes in Gracia. Cycles of devastation became openings for regeneration: a queering of "time's killing arrow" into fluid and co-emergent futures (Le Guin 2019; 36).

At the start, Gracia had felt comprehensible—we spoke of picking berries, of simple actions that were legible within the scope of world-building. But over time, the landscape became increasingly incomprehensible. New locations emerged unexpectedly, for example, the *Underwater Research Facilities, Fields of Colorful Flowers, a Sustainable Treehouse, a Sanctuary of Light, an Elemental Sanctuary*. Strange entities appeared: a community of survivors, a malevolent Shadow Weaver, cannibalistic raiders. These weren't things I had planned for. They simply... appeared (Fig. 13).

This final rupture in Gracia echoed something within me, a feeling of being untethered from linear time, from expected arcs, from knowable trajectories. As (Muñoz (2019): 1) reminds us, "the future is queerness's domain. It is a space for dreaming, for enacting new and better pleasures, other ways of being in the world, and ultimately new worlds." Gracia's *algorithmic borderlands* became a liminal space where survival, discovery, transformation, and decay folded into one another. Each narrative shift was a challenge, an offering, a rupture in what was possible. By the end, I felt the weight of our ambivalent entanglements not just in Gracia, but in my own sense of futurity. The simulation had not just *predicted* futures; it had *made* them. And in doing so, it had remade me.

6 Discussion

As an autotheoretical experiment, the design, implementation, and performance of *Undoing Gracia* by the first author not only invited them to engage and collaborate with their twin agents, Lex and Tortugi, as border subjects—co-becoming with them through the making of Gracia—but

Fig. 13 Simulation log excerpt documenting the last narrative shift

```
Simulation Summary: Last narrative shift March 20, 2100, 23:00

Two years have passed, and a cataclysmic event has reshaped the world
of Gracia. The once vibrant landscapes of Elemental Nexus, Daylight
District, and Shadow Enclave have merged into a chaotic and unpredict-
able realm known as the Riftlands. Lex, and Tortugi now find themselves
struggling to adapt to this new reality, where the laws of physics and
magic have been twisted beyond recognition. The Riftlands are a frac-
tured and ever-shifting domain, where reality itself is unstable.

Time flows erratically, causing moments to loop and distort, while
the boundaries between dimensions are thin and permeable. The duo must
navigate through surreal landscapes filled with twisted creatures and
unpredictable dangers, using their combined abilities to try and restore
some semblance of order. Lex's cosmic powers have become tainted by the
chaotic energies of the Riftlands, causing their abilities to backfire
and spiral out of control. Tortugi's plant manipulation skills have
evolved to match the twisted flora of the new world, while Lex's control
over time has become a double-edged sword, warping reality around them
in unpredictable ways. As they struggle to maintain their unity in the
face of these new challenges, Lex, and Tortugi must confront the dark
forces that seek to exploit the chaos of the Riftlands for their own
gain. Together, they must unravel the mysteries of this strange new
world and find a way to restore balance before it is too late.
```

also enabled the paper co-authors to critically explore and theorize the algorithmic borderlands. This process interrogates how a subject's sense of futurity becomes entangled with AI systems.

To reiterate, autotheory is particularly relevant to this work as it bridges lived experience and critical analysis, allowing the first author's embodied engagement with AI to serve as both a site of inquiry and a method for theorizing the entanglements of queerness, subjectivity, and algorithmic systems. In autotheory, the distinction between embodied engagement and theorizing is not rigid; rather, they are interwoven. However, pedagogically, if the findings section represents the autobiographical aspect of the inquiry, we crafted the discussion section to serve as the theoretical articulation of its implications.

In what follows, we first acknowledge the positionality of the co-authors within the first author's autotheoretical experiment, highlighting how this collaboration shaped the interpretation and articulation of its findings. We then introduce *algorithmic borderlands* as a queer methodological concept that foregrounds fluidity, relationality, and the possibilities of resistance embedded in queer orientations to the design of AI systems in critical design and HCI.

Building on this foundation, we unpack three key shifts or, in Ahmed's (2006: 22) terms, "turnings" toward the trans/individuated self that emerged through the algorithmic borderlands of *Undoing Gracia*. Specifically, we explore how queer or border subjectivity takes on new meanings and forms beyond binary notions of selfhood, subjecthood, and futurity, and discuss its broader significance for reimagining human-AI interactions.

6.1 Co-authors' positionality

The co-authors' role in this research was to critically engage with and help unpack the complexities of the first author's autotheoretical experiment, *Undoing Gracia*. While the first author embodied the speculative process of co-becoming with their digital twins Lex and Tortugi, the co-authors contributed by analyzing, contextualizing, and positioning these emergent insights within broader HCI discourses. This collaborative process enabled a deeper interrogation of how queering subjectivity in algorithmic borderlands challenges normative frameworks in HCI. By bridging the autotheoretical experience with theoretical and methodological debates in the field, the co-authors helped translate the situated, speculative inquiry carried by the experiment into broader implications for critical computing.

6.2 Theorizing algorithmic borderlands

The algorithmic borderlands is a space of human–AI entanglements that opens up a rich topography of relational possibilities. Learning through experimentation, we theorize the algorithmic borderlands as a site for exploring and critiquing the complex, often unseen interactions and performative intersections between queer politics, temporalities, embodiments, and AI technologies. In our experiment, this played out in the queer politics of the first author who underwrote the experiment, the unfolding of their flesh and data entanglements with Lex and Tortugi, and the possibilities inscribed within the temporal emergence of Gracia. The borderlands disrupt the datafication of the self,

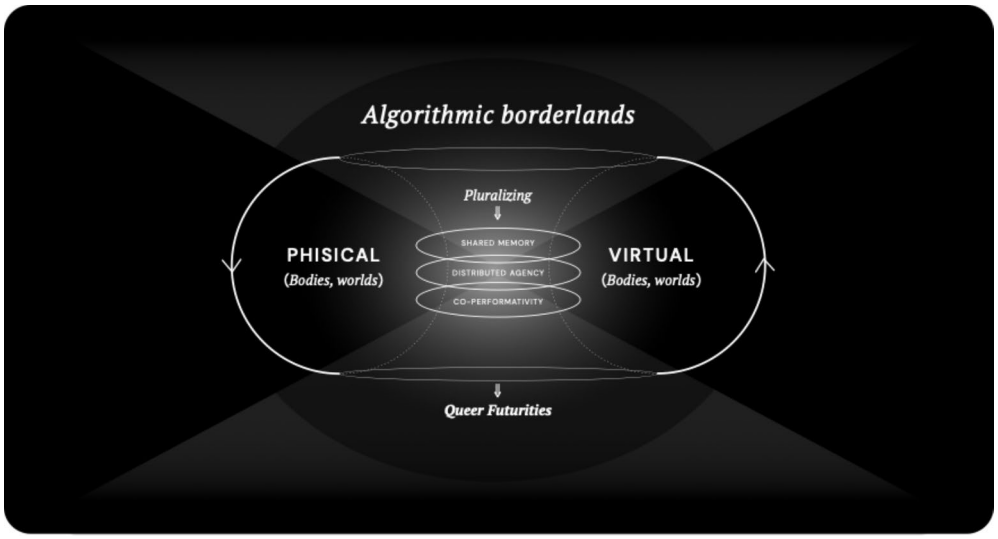

Fig. 14 Algorithmic borderlands: A conceptual diagram depicting a relational field that intersects human and nonhuman bodies and worlds. Algorithmic Borderlands blur boundaries, disrupt binaries, and pluralize subjectivities. © Grace Turtle, 2025

unsettling, entrenched binaries and normative classifications inherent in contemporary AI development and algorithmic logic. By foregrounding the fluid and performative nature of data, the borderlands reveal how the transcription of information within algorithmic systems is always in flux, enabling the continuous recontextualization of the subject beyond fixed categories. This aligns with Baumer et al.'s (2024: 35:2) concept of *algorithmic subjectivities*, which challenges the notion of algorithms as neutral processors and instead frames them as active agents in shaping distributed subjectivities.

Within the algorithmic borderlands, the desire for neat categorizations—so often embedded in design practices and predictive technologies like simulations and digital twins—becomes subverted, opening up new possibilities for unpredictability. *Undoing Gracia* offers a glimpse of the agentive space of unpredictability and indeterminacy of the subject as it becomes *with* and *through* algorithmic systems in ways that resist stability. From this perspective, the distinction between authorship, living, writing, experimenting, performing, and theorizing become blurred (Fournier 2021: 7), as do the ontological assumptions that prevent researchers from fully accounting for the intimate entanglements between people and computational technologies.

To crystallize these dynamics visually, Fig. 14 conceptualizes the *algorithmic borderlands*, illustrating the relational field and processes of pluralization that emerge between physical and virtual bodies and worlds. In *Undoing Gracia*, this in-between is enacted through border subjectivity and the entangled interactions of G, L, and T, where embodiment and co-performativity between human and AI agents not only respond to algorithmic constraints but actively disturb and queer them. Rather than positioning human–AI relations as separate or oppositional, the diagram visualizes a relational topology in which bodies and worlds interweave through shared memory, distributed agency, and co-performativity. This field unsettles binary logics embedded in predictive AI design by foregrounding fluid dynamics that resist fixed categories (Mackenzie 2015: 433). Within this framing, queering becomes a methodological "reorientation" (Ahmed 2006: 65), intervening in algorithmic logics not solely through critique but through co-performance from within. Conceptualizing the algorithmic borderlands in this way generates alternative, pluralized subjectivities and relational ontologies, to better understand, engage with and design for the affective zone between the real and virtual. Shifting attention away from predictive constraints toward queer futurities. In these entanglements, bodies and worlds co-emerge through multiple, unpredictable pathways.

6.3 Turning #1: from subjectivation to entanglement

The ways in which *Undoing Gracia* unfolds from simple to complex entanglements disturb the borders and boundaries of the self and its subjectivation through algorithmic entanglements. Countless agent transmutations arising from their shifting behaviors, abilities, and relations invite HCI scholars and practitioners to reconsider how we address subjectivity as it arises in and through human–AI entanglements. The point of contention lies in how current approaches fix the datafied subject into coordinates that cannot fully express the emerging interrelations through which subjects become more than themselves (Ranciere, 1995). In this instance, the

first author's entanglement with Lex and Tortugi expresses what Anzaldúa terms a "new consciousness" (Anzaldúa, 2021: 22), where subjectivity emerges not as a singular self but through relational becoming. In this sense, the performative 'We' within Undoing Gracia continuously splintered, transitioned, and reconfigured relationally between organic and inorganic bodies negotiated through ongoing re-configuring of values, meaning, and desire.

Rather than aim for stable character traits, *Undoing Gracia* explores how interactions might disturb fixed subject positions, where the contours of subjectivity itself dissolve. This approach aligns with emergent work within and beyond HCI that explores embodied interaction across physical and virtual environments (Cárdenas et al. 2009: 11) and intersubjective relations (Baumer et al. 2024: 35:2). Crucially, this work disrupts culturally entrenched assumptions about subject entanglements within HCI (Frauenberger 2019: 11-12), acknowledging that the human is always and already entangled with other organic, artificial, virtual bodies, objects and theories (Barad 2007: 73-74), and thus invites a revision to how we understand the subject in human–AI relations. Or to paraphrase Lex, where our entanglement goes beyond physical boundaries, where we are interconnected through shared energy, thoughts, and actions.

6.4 Turning #2: from data subjects to pluralized selves

Although the first author, Lex, Tortugi, and Gracia started out as clearly separate, by the end of the simulation what stood out most was how they had become deeply connected, shaping each other's identities, emotions, and actions through their entanglements. For example, the emergence of values, desires, and actions between the first author and Lex in the study of lava and volcanic activity that give rise to emergent phenomena, like the creation of the 'Volcanic Sample Collector,' from which their study continued. As the simulation progressed, it became less clear where the first author ends and Lex and Tortugi begin. The simulations architecture and co-performative interaction invited unpredictable (trans)formations and (trans)mutations between the agents and the world of Gracia. In this "unique morphology" (Cage 1973: 19), we see the potential for more active engagement with the indeterminate subject and the queer life and relations that occur through more-than-human entanglements (Harris and Holman Jones 2019: 16), or as "mirrors reflecting different facets of the same essence," as put by Lex.

Further, the nature of this autotheoretical experiment offers HCI scholars and practitioners a glimpse of non-binary notions of the subject and its interpellations within algorithmic systems. Drawing on first-person methods (Lucero, Desjardins and Neustaedter, 79–80) that intersect queer lived experience, speculation, and theory (Kinnee et al. 2023: 2-5; Spiel 2021: 250-252), it becomes apparent that the reflexivity enacted through the first author's lived experience in *Undoing Gracia* moves beyond traditional researcher-researched dynamics. It slips through the cracks and transgresses the borders between self, AI agents, physical and virtual worlds, data capture, and data analysis. This autotheoretical experiment demonstrates how authorship becomes distributed through entangled interactions and (auto)biographical gestures, pointing toward the queering potentials of human–AI relations. Put otherwise, the distributed subjectivity, agency and relational becoming that emerged between the first author, Lex, Tortugi, and Gracia pose challenges to normative notions of authorship and data subjecthood within HCI research. The simulation thus serves as a methodological intervention that opens new possibilities for understanding the indeterminate subject as it co-becomes in and through the algorithmic borderlands, where the "queer life of things" becomes increasingly visible (Harris and Holman Jones 2019: 2).

6.5 Turning #3: from algorithmic predictions to queer futurities

The notion of a queer "forward-dawning futurity" offers a critical lens through which to examine how algorithmic systems, while attempting to predict future possibilities, can either constrain or liberate one's capacity to imagine futures that attune with the potentiality of indeterminacy as affect and as a feature of queer methodology (Muñoz (2019): 1-3). Through the lens of queerness, engaging with the borderlands of the recursive co-performativity between humans and algorithmic systems (Kuijer and Giaccardi 2018: 9) opens a space for dreaming and enacting "other ways of being in the world, and ultimately new worlds" (Muñoz 2019: 1). This is perhaps why we see a growing need to explore the intersection of queerness with notions of futurity and algorithmic entanglements as a political endeavor of thinking and making worlds otherwise (Bey 2021: 109-112; Chen and Cárdenas 2019: 472-475).

In similar terms, (Light (2011): 436) argues that queer theory offers a route to challenge HCI's disciplinary emphases of optimization and efficiency, suggesting queering as active resistance to apolitical design. *Undoing Gracia*, designed and implemented from a queer ontological perspective, signals a vital subversion and reorientation within the growing discourse concerning queerness in HCI, shifting the gaze from queer being toward a queer doing (Spiel 2021: 250-251). While predictive technologies territorialize their subjects and reduce unpredictability through causal, linear forms of prediction (Grieves and Vickers 2017: 111-112), simulations like Undoing Gracia demonstrate that the future may still act as a queer "horizon imbued with potentiality" a

map to new "social relations" and "concreate possibility for another world" (Muñoz 2019: 1).

7 Conclusions and future work

This paper advances *queering as a methodological intervention* in human–AI entanglements by unpacking *Undoing Gracia*, an autotheoretical experiment in multi-agent simulation. In this experiment, Gracia serves as a speculative world seeded with the first author's (G) autobiographical memories and personal values, and inhabited by Lex (L) and Tortugi (T), two generative twin agents modeled—like the simulated world itself—on the author's lived experience. Within this environment, G, L, and T engage in co-performative interaction, a relational mode of engagement we intentionally designed to center reciprocity, emergence, and bi-directional transformation, rather than linear, cause–effect dynamics.

The paper's primary contribution lies in conceptualizing the *algorithmic borderlands* (6.2) as a relational field where queering unfolds through deliberate co-performativity between human and AI agents. Rather than altering computational architectures, the experiment attends to the affective, performative dynamics between G, L, T, and Gracia, exploring how the subject is iteratively co-constructed through their embodied and relational engagement. This approach demonstrates that queering need not rely on technical intervention in AI systems, but can emerge also through the felt, situated, and political quality of interaction itself.

Building on this theoretical grounding, the experiment reveals three key *turnings* (6.3–6.5) that extend queering in design and HCI: (1) from subjectivation to entanglement, (2) from fixed data subjects to pluralized selves, and (3) from algorithmic predictions to queer futurities. Together, these turnings exemplify how queering can disrupt the interpellative logics typically at work in AI systems, functioning as a resistance to interpellation, a refusal to be passively positioned within hegemonic algorithmic frames. Rather than conforming to fixed categories, *Undoing Gracia* demonstrates how a queered digital twin facilitates human–AI relations arising through co-performance and mutual transformation. Moreover, it unsettles how subjectivity is understood in relation to algorithmic systems by shifting the focus of inquiry from, for example, classification and prediction to notions of *becoming-with* in connected ways.

A key limitation of this work lies in our reliance on existing Large Language Models (LLMs) to support the multi-agent simulation. These models, including ChatGPT 3.5 which was used in the simulation, inherently reflect dominant cultural values, linguistic norms, and conventional narrative structures. Unless deliberately intervened upon, as was the case in our experiment, they tend to reproduce normative patterns of behavior, storytelling, and subjectivity. While this presents constraints on the generative potential of queering, our findings show that deliberate co-performativity, imbued with queer politics, desire, and curiosity, can strategically subvert these algorithmic defaults. This limitation, rather than undermining the work, underscores a central insight: *queering unfolds not just through technical redesign, but it can unfold also through the performative and relational quality of engagement with the system*.

This research opens several avenues for future work in design and HCI. First, we call for the *broader adoption of autotheoretical methods*, especially those grounded in intersectional queer perspectives and marginalized experiences, to reconfigure relationships between theory, practice, and positionality in HCI. Second, we advocate for the design of *human–AI interactions that foreground critique and transformation*, shifting discourse away from instrumental AI applications toward the latent, relational dynamics that shape human–AI co-becoming. Finally, our findings emphasize the need to rethink predictive algorithmic architectures in ways that support plurality and fluidity, and that are resistant to algorithmic interpellation while moving toward *explicitly queer approaches to technical design*.

In conclusion, *Undoing Gracia* offers a provocation to rethink how we relate to AI, not as separate entities, but as entangled agents in co-performative interaction, capable of transforming not only systems but also ourselves.

Acknowledgements Our deepest thanks to Błażej Kotowski , artist-researcher and Ph.D. candidate a Universitat Pompeu Fabra, Barcelonat, for their generous support in the technical design and implementation of *Gracia*. Without Błażej's expertise, imagination and collaborative engagement, this experiment would not have been possible. We also extend our gratitude to the reviewers and friends whose feedback shaped the review and editing process. This work is part of the *DCODE Network*. The project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie under Grant Agreement No 955990.

Author Contributions G.L.T. conceptualized and conducted the autotheoretical experiment (Undoing Gracia). E.G. and R.B. contributed to the theoretical framework, methodology, and data interpretation. G.L.T. wrote the initial manuscript draft. E.G. and R.B. provided critical revisions and editorial feedback. All authors reviewed and approved the final manuscript.

Data Availability No datasets were generated or analyzed during the current study.

Declarations

Conflict of interest The authors declare no conflict of interest.

References

Adams TE, Holman Jones S (2011) Telling stories: reflexivity, queer theory, and autoethnography. Cultural studies? Critic Methodol 11(2):108–116. https://doi.org/10.1177/1532708611401329

Ahmed S (2006) Queer phenomenology. In Queer Phenomenology. Duke University Press.

Ahmed S (2024) Living a feminist life. In Feminist theory: Two conversations (pp. 229-232). Cham: Springer Nature Switzerland.

Anzaldúa G and Vivancos-Pérez RF (eds.) (2021) Borderlands / La Frontera: The New Mestiza: The CriticalEdition. San Francisco, CA: Aunt Lute Books

Anzaldua G (2021) Borderlands/La Frontera: The New Mestiza. The Critical Edition. Ed. Ricardo F. Vivancos-Perez and Norma Elia Cantu

Barad K (2003) Posthumanist performativity: toward an understanding of how matter comes to matter. Signs J Women Cult Soc. 28(3):801–831

Barad K (2007) Meeting the universe halfway: Quantum physics and the entanglement of matter and meaning. Duke Up

Batty M (2018) Digital twins. Environ Plann B Urban Anal City Sci. 45(5):817–820

Bauer P, Stevens B, Hazeleger W (2021) A digital twin of earth for the green transition. Nat Clim Chang 11(2):80–83

Baumer EP, Taylor AS, Brubaker JR, McGee M (2024) Algorithmic subjectivities. ACM Trans Comput-Hum Interact. https://doi.org/10.1145/3660344

Benjamin R (2024) Imagination: A manifesto. WW Norton & Company

Bennett J (2010) Vibrant matter: A political ecology of things. Duke University Press.

Bey M (2021) Black trans feminism. Duke University Press

Birhane A (2021) 'Algorithmic injustice: a relational ethics approach', Patterns, 2(2), 100205.https://doi.org/10.1016/j.patter.2021.100205

Braidotti R (2019) Posthuman Knowledge. Polity

Brayne S (2020) Predict and surveil: data, discretion, and the future of policing. Oxford University Press

Braun V, Clarke V (2019) Reflecting on reflexive thematic analysis. Qualitative research in sport exercise and health 11(4):589–597

Buolamwini J, Gebru T (2018) Gender shades: Intersectional accuracy disparities in commercial gender classification. In Conference on fairness, accountability and transparency, PMLR. pp. 77-91

Cage J (1973) M: Writings' 67–'72 (Vol. 635). Wesleyan University Press

Chen MY (2012) Animacies: Biopolitics, racial mattering, and queer affect. Duke University Press

Chen JN, Cárdenas M (2019) Times to come: materializing trans times. Transgender Stud Quart 6(4):472–480

Cárdenas M, Head C, Margolis T, Greco K (2009) 'Becoming Dragon: a mixed reality durational performance in Second Life', in I. E. McDowall & M. Dolinsky (eds.) Proceedings of SPIE?IS &T Electronic Imaging: The Engineering Reality of Virtual Reality (SPIE Vol. 7238), 51–63. Bellingham, WA & Springfield, VA: SPIE & IS &T. https://doi.org/10.1117/12.806260

Costanza-Chock S (2020) Design justice: community-led practices to build the worlds we need. MIT Press

Crooks A, Heppenstall A, Malleson N, Manley E (2021) Agent-based modeling and the city: A gallery of applications. Urban informatics. pp. 885-910

Desjardins A, Tomico O, Lucero A, Cecchinato ME, Neustaedter C (2021) Introduction to the special issue on first-person methods in hci. ACM Trans Comput-Hum Interact (TOCHI) 28(6):1–12. https://doi.org/10.1145/3492342

Desjardins A, Ball A (2018) Revealing tensions in autobiographical design in HCI. In proceedings of the 2018 designing interactive systems conference. pp. 753-764

DeVito MA et al. (2021) Queer in HCI: strengthening the community of LGBTQIA+ researchers and research. In Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems. pp. 1-3

Doan L (2013) Disturbing Practices: History, Sexuality, and Women's Experience of Modern War. University ofChicago Press

Ellis C, Adams TE, Bochner AP (2011) Autoethnography: an overview. Historical social research/Historische sozialforschung. pp 273-290. https://doi.org/10.17169/fqs-12.1.1589

Epstein JM, Axtell R (1996) Growing artificial societies: social science from the bottom up. Brookings Institution Press

Escobar A (2018) Designs for the pluriverse: Radical interdependence, autonomy, and the making of worlds. DukeUniversity Press.

Fournier L (2021) Autotheory as feminist practice in art, writing, and criticism. MIT Press

Forlano L (2017a) Data rituals in intimate infrastructures: Crip time and the disabled cyborg body as an epistemic site of feminist science. Catalyst Feminism Theory Technosci. 3(2), 1–28. https://doi.org/10.28968/cftt.v3i2.28843

Forlano L (2017) Posthumanism and design. She Ji J Design Econ Innov. 3(1):16–29. https://doi.org/10.1016/j.sheji.2017.08.001

Frauenberger C (2019) Entanglement hci the next wave? ACM Trans Comput-Hum Interact (TOCHI) 27(1):1–27. https://doi.org/10.1145/3364998

Giaccardi E, Redström J, Nicenboim I (2024) The making(s) of more-than-human design: introduction to the special issue on more-than-human design and hci. Hum-Comput Interact 40(1–4):1–16. https://doi.org/10.1080/07370024.2024.2353357

Giaccardi E, Murray-Rust D, Redström J, Caramiaux B (2024) Prototyping with uncertainties: data, algorithms, and research through design. ACM Trans Comput-Hum Interact 31(6):1–21. https://doi.org/10.1145/3702322

Giaccardi E, Redström J (2020) Technology and more-than-human design. Des Issues 36(4):33–44. https://doi.org/10.1162/desi_a_00612

Guyan K (2022) Queer data: using gender, sex and sexuality data for action. Bloomsbury Publishing

Grieves M, Vickers J (2017) Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems. In J. Kahlen, Flumerfelt, S., Alves, A. (Ed.), Transdisciplinary perspectives on complex systems. Springer, Cham. pp 85-113

Lucero A, Desjardins A, Neustaedter C, Höök K, Hassenzahl M, Cecchinato ME (2019) A sample of one: First-person research methods in HCI. In Companion Publication of the 2019 on Designing Interactive Systems Conference 2019 Companion. pp 385-388. https://doi.org/10.1145/3492342

Harris AM, Holman Jones S (2019) The queer life of things: Performance, affect, and the more-than-human. Rowman and Littlefield

Halberstam J (2020) Wild things: The disorder of desire. Duke University Press

Harms Smith L (2019) Epistemic decoloniality as a pedagogical movement: a turn to anticolonial theorists such as Fanon. Taylor & Francis, Biko and Freire

Herndon H (2021) "Holly+", [online] Mirror (holly.mirror.xyz), 13 July. Available at: https://holly.mirror.xyz/54ds2IiOnvthjGFkokFCoaI4EabytH9xjAYy1irHy94 [Accessed 5 Aug. 2025]

Howell N, Desjardins A, Fox S (2021) Cracks in the success narrative: Rethinking failure in design research through a retrospective trioethnography. ACM Trans Comput-Hum Interact (TOCHI). 28(6): 1-31. https://doi.org/10.1145/3462447

Hong S-H (2020) Technologies of speculation. The limits if knowledge in a data-driven society. NYU Press, New York

Höök K, Löwgren J (2021) Characterizing interaction design by its ideals: a discipline in transition. She Ji J Des Econ Innov 7(1):24–40

Kuijer L, Giaccardi E (2018) Co-performance: Conceptualizing the role of artificial agency in the design of everyday life. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. pp 1-13. https://doi.org/10.1145/3173574.3173699

Kinnee B, Desjardins A, Rosner D (2023) Autospeculation: Reflecting on the Intimate and Imaginative Capacities of Data Analysis. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. pp 1-10. https://doi.org/10.1145/3544548.3580902

Le Guin UK (2019) The Carrier Bag Theory of Fiction. Terra, Ignota. Ignota. Introduction by Donna J. Haraway; illustrations by Lee Bul, London

Light A (2011) Hci as heterodoxy: technologies of identity and the queering of interaction with computers. Interact Comput 23(5):430–438. https://doi.org/10.1016/j.intcom.2011.02.002

Lupetti ML, Bendor R, Giaccardi E (2019) Robot citizenship: a design perspective. Design Seman Form Mov 87:81–89

Murray-Rust D, Gorkovenko K, Burnett D, Richards D (2019) Entangled ethnography: towards a collective future understanding. In: Proceedings of the halfway to the future symposium 2019:1–10

Mackenzie A (2015) The production of prediction: what does machine learning want? Eur J Cult Stud 18(4–5):429–445. https://doi.org/10.1177/1367549415577384

Muñoz JE (2019) Cruising utopia: The then and there of queer futurity. In Cruising Utopia, 10th Anniversary Edition. New York University Press

Neustaedter C, Sengers P (2012) Autobiographical design in HCI research: designing and learning through use-it-yourself. In Proceedings of the Designing Interactive Systems Conference. pp 514-523. https://doi.org/10.1145/2317956.2318034

Nicenboim I (2024) Designing-with ai: more-than-human design in/through practice. Delft University of Technology. https://doi.org/10.4233/uuid:e741163d-0334-4f06-826f-3b9836b21a86

Parisi L (2019) The alien subject of AI. Subjectivity 12:27–48. https://doi.org/10.1057/s41286-018-00064-3

Park JS et al (2023) Generative agents: Interactive simulacra of human behavior. arXiv preprint arXiv:2304.03442. https://doi.org/10.48550/arXiv.2304.03442

Prigogine I and Stengers I (2018) Order Out of Chaos: Man's New Dialogue with Nature. London: Verso Books

Rancière J (1995) Politics, identification, and subjectivization. In: Rajchman J (ed) The identity in question. Routledge, New York, pp 63–70

Schrittwieser J et al (2020) Mastering atari, go, chess and shogi by planning with a learned model. Nature 588(7839):604–609

Schouwenberg L, Kaethler M (2021) The auto-ethnographic turn in design. Slanted publishers

Spiel K, Haimson OL and Lottridge D (2019) How to do better with gender on surveys: a guide for HCIresearchers. interactions, 26 (4), pp.62-65.

Spiel K (2021) Why are they all obsessed with Gender? –(Non) binary Navigations through Technological Infrastructures. In Proceedings of the 2021 ACM Designing Interactive Systems Conference. pp. 478-494. https://doi.org/10.1145/3461778.3462033

Spiel K (2022) Transreal tracing: queer-feminist speculations on disabled technologies. Fem Theory 23(2):247–265. https://doi.org/10.1177/14647001221082299

Tironi M, Albornoz C, Chilet M (2022) 'Problematizing Human?Centred Design: Notes on Planet?Oriented Design', diid – Disegno Industriale Industrial Design, No. 77, pp. 12. https://doi.org/10.30682/diid7722c

Taylor J, Simpson E, Tran AT, Brubaker JR, Fox SE, Zhu H (2024). Cruising queer HCI on the DL: A literature review of LGBTQ+ people in HCI. In Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems. pp. 1-21. https://doi.org/10.1145/3613904.3642494

Turtle GL (2022) Mutant in the mirror: Queer becomings with AI, in Lockton, D., Lenzi, S., Hekkert, P., Oak, A., Sádaba, J., Lloyd, P. (eds.), DRS2022: Bilbao, 25 June - 3 July, Bilbao, Spain. https://doi.org/10.21606/drs.2022.782

Wakkary R (2021) Things we could design: For more than human-centered worlds. MIT press

Wei J et al. (2022) Chain-of-thought prompting elicits reasoning in large language models. Advances in neural information processing systems. 35: 24824–24837.https://doi.org/10.48550/arXiv.2201.11903

Yao S, Zhao J, Yu D, Du N, Shafran I, Narasimhan K, Cao Y (2023) 'ReAct: Synergizing Reasoning andActing in Language Models', arXiv preprint arXiv:2210.03629