

Proposal: Grey Media: Gaslighting, Post-Truth, AI Deception

Prof Lisa Blackman, Department of Media, Communications and Cultural Studies,
Goldsmiths, University of London, UK.

l.blackman@gold.ac.uk

Rationale: Why Punctum Books?

Grey Media is a genre-defying and genre-bending book, a structure that reflects the wild convoluting and confabulating realities that are staged throughout the writing. The book **intervenes** within a range of contemporary debates and controversies which are problematizing, exploring, unsettling, and reimagining the nature and status of deception as it emerges as an object in the context of AI, information warfare, post-truth politics and the gaslighting associated with psychological and emotional abuse. The project is **distinctive** in its radical interdisciplinarity or even *undisciplinarity*. It moves across the sciences and humanities, the popular and the political, and the actual and the compositional, focusing on examples which are designed to engage a more general reader interested in the “power of possibility” and its close relationship to deception. The book navigates and weaves together interdisciplinary debates that cross cultural and media studies, affect studies, *critical* mental health, psychology, and disability studies, speculative philosophies, feminism and queer theories. The project is shaped in a conjuncture where understanding processes and practices of deception has been increasingly placed at the centre of contemporary debates about power, networks, information, truth, reason, and emotion, and what is constituted as social pathology in a contemporary conjuncture. This includes what registers as odd, strange, weird, bizarre, and even stranger than fiction.

The project takes popular notions of gaslighting as a springboard to explore contemporary modes of power that operate through forms of storytelling that play with indeterminacies of sense, feeling and perception. They challenge and unsettle liberal moral norms, weaponize the dilemmatic nature of affective and social life, and circulate within relations of distrust, disbelief, uncertainty, and confusion, creating zones of indescribability, opacity, and discordance. The modes of storytelling are associated with both creativity *and* disorder, representing the paradoxical nature of confabulation, a form of emotional reasoning that challenges distinctions between the rational and the irrational, the legible and the illegible, the moral and the immoral, intent and opportunity. The book maps modes of deceptive storytelling that have common structures of communication and unfeeling, whilst staging and performing some of the strange and illogical couplings of feeling-rules and actions, which are part of their operations.

The book develops a novel approach to the commonalities explored throughout the chapters, combining historical and archival analysis, and a method of speculative storytelling told by a fictional character called Ada. Ada is a queer figure, an unstable narrator who composes scenes for the reader that are based on a *weak* ethnography (developed from the concept of weak fiction) of the lived experiences of targets of non-physical abuse, and their mediation via blogs, Tik Tok videos, podcasts, and social media forums. The project engages the reader in counterfactual scenarios that stage the subject-matter of the book, composing affect-laden scenes that play with the boundaries of fiction and nonfiction, self and other.

These scenes draw from a variety of data sources, news sources, journalistic accounts, commentaries by pundits, stories propelled by social influencers, as well as research and the experiences of dissimulation and deception by those with direct lived experience¹.

The book moves beyond the conventional monograph, combining different modes of writing, which are attuned to and resonate in contexts where the subject matter under analysis can only be made and spoken within very specific contexts of production, reception and even survival. The project is highly suitable for a publisher such as Punctum, with its commitment to books that create and shape transdisciplinary conversations that are often siloed in particular disciplinary fields of research with distinct readerships. The book is speculative, whilst engaging with what haunts creative speculative thought in the context of abuses of power. It centres confabulation within a biopolitics of convolution that operates through battles for the control of narrative, multiplying what could be through “what if” scenarios of both sense and nonsense. I believe the book represents the ethos and spirit of Punctum Books, with its commitment to cutting across and blurring disciplinary boundaries of knowledge and practice, and form and content.

Chapter Breakdown

Introduction –Confabulating and Convoluting Realities sets the scene for the book by mapping and staging some of the illogical and strange couplings of feeling-rules and actions which form the subject matter of *Grey Media*. The introduction provides an overview and various entry points into the project for different readerships, including readers interested in AI Deception, post-truth, and the gaslighting of emotional and psychological abuse. The book is situated in a conjuncture where I argue that deception has become an attractor for a range of issues: framed as a set of norms, strategies, a moral status, and even programmable as part of “malicious influence campaigns”ⁱ. The project is shaped in a context where understanding processes and practices of deception has been increasingly placed at the centre of contemporary debates about power, networks, information, truth, reason, emotion, and what is constituted as social pathology, including what registers as odd, strange, weird, bizarre, and even stranger than fiction.

Chapter one, *Gaslighting-as-Media*, develops the novel shift in the book from approaching gaslighting as an ensemble of psychological techniques of manipulation and control to gaslighting as a form of media. It starts with a series of questions designed to engage the reader in a counterfactual scenario to provide a hook and way into some of the material central to the arguments developed throughout the book. The chapter provides an overview and development of the approach to gaslighting-as- media and its importance for analyses of power that link political and emotional abuse. The chapter situates the book within the context of battles for the control of narrative across war, the military, soft power, non-physical abuse, and politics, outlining its indebtedness to Jackie Orr’s work on PSYCHOpower

¹ This includes data drawn from a three-year study of participants stories across social media forums who understand their experiences through the concept of narcissistic abuse. Narcissistic abuse is a term that has been popularized in different ways, described as “all the rage” across social media. It has become a stand in for both gaslighting and the social figure of the gaslighter.

(2009), Matthew Fuller and Andrew Goffey's (2012) work on Evil Media, and the late American and cultural studies theorist, Lauren Berlant's (2005) insightful analysis of disbelief as a political emotion. It outlines some of the key concepts central to the book (confabulation and convolution) which are put to work in analyses of post-truth, AI Deception, and the experiences of people who have come together within social media forums to narrativize and re-story their experiences through the concept of narcissistic abuse.

Chapter 2, *Nonsense* turns more explicitly to the psychological sciences and the histories of controversy and epistemic uncertainty that surround a range of phenomenon and practices that have a kinship to experiences of lying, dissimulation, and deception that we are now more likely to term gaslighting. Gaslighting is one more example of forms of emotional reasoning that have been considered nonsense and "odd phenomenon" for the challenges they pose to the bounded universal psychological subject (Eysenck, 1956). At the same time, they have furnished the humanities with ways of understanding mediation that confound distinctions between subject and object, human and technical, fact and fiction, past and present, the rational and the irrational and the anomalous and the ordinary. The chapter explores some of these issues by exploring the relationship of psychological understandings of gaslighting to the concept of confabulation, a form of emotional reasoning that is understood as necessary for creativity, whilst at the same time it carries the potential to err, to create faulty perceptions, and even counterfactual scenarios that play with the power of possibility. The potential to err is contained within psychology through the concept of disorder, aligned to an archive of disordered brains, erring cognitive processes, and deficits in emotion associated with the figure of the narcissist and the psychopath.

The chapter explores how confabulation challenges what appears under the sign of normal psychological health, showing how cognition or reason *feel*, an argument that has been incisively made by Donovan Schaefer (2022) in relation to the tendency of science, and in this context the psychological sciences, to separate thinking from feeling, and affect from cognition. The chapter draws out some of the issues by bringing together work on the use of confabulation as a form of creative imagination and speculative method across the humanities, with research which explores the mediality of how it feels to be subject to the more malign aspects of confabulation within the context of power and abuse. This includes the experience of being trolled, as recounted by the educational scholar Jessica Ringrose in the context of her research challenging mediated misogyny, the experience of malevolent technologies of coercive control in the context of psychological abuse, and an ethnographic analysis of gaslighting within non-government organisations, as recounted by employees of NGO's. The chapter begins the work of expanding psychological understandings of gaslighting laying the groundwork for later chapters analysing gaslighting as a form of media that normalizes both the creative and more malign aspects of confabulation as a form of counterfactual reasoning.

Chapter 3, *Information Warfare*, develops analytic insights from what I call a *weak* ethnography of experiences of people who have gathered under the sign of narcissistic abuse to articulate and attempt to make sense of their experiences. I approach these groups as experts-by-experience, providing important insights into more public processes and practices of communicative control and semiotic combat based on deception and fraudulent

perceptions. I borrow the term “weak”ⁱⁱ ethnography from the concept of the weak novel (see Dimock, 2020), which works with fiction as strategy, rather than entertainment. The concept of a weak ethnography, in the context of non-physical abuse, acknowledges that accounts will be replete with digressions, will be incomplete and nonlinear, where there are non-narrative elements that remain unspoken, where subjects speak from a place of vulnerability, and where there will be gaps and silences.

Weak ethnographies play with the boundaries of fiction and non-fiction, truth and falsehood, self and other. They resonate in contexts where the subject matter under analysis can only be made and spoken within very specific contexts of production, reception and even survival. This is primarily due to the stigma, danger, and risk that articulations of experiences might pose to research collaborators and subjects. The ethnography is based on the author’s participation over a three-year period with many people who have joined open and closed social media forums dedicated to the problem of narcissistic abuse, and through my engagement with blogs, podcasts, memoirs, Tik Tok videos, hashtags, and other practices of wi(th)nessing, which conjoin the figure of the narcissist, a stand-in for emotionally abusive manipulation, with a variety of tactics that gaslighting becomes a shorthand for. The chapter aims to stage the affectivity of the signs and symbols associated with the forms of deception and fraudulent perception that are a key feature of “narcissistic abuse”, and the dynamics and relationality of the information flows that are part of disorienting and distorting exchanges that we might describe as gaslighting.

Chapter 4, *Disbelief as Political Emotion*, explores the affective and emotional experiences of convoluting forms of power and their obfuscation of the perception of deception in contexts of personal and political confabulation. The chapter foregrounds the political and performative work that disbelief does, as forms of political and social misdirection common to non-physical abuse and post-truth politics. Within the context of post-truth and Trumpism, for example, disbelief as a political emotion has been recognised as an important part of a set of political strategies designed to create the setting for convolution and discordance. Rather than cognitive dissonance, a psychological term often used to describe the affective and embodied experience of a chaotic and volatile communication-political environment, I explore the registering of discordance through what Lauren Berlant (2005) called affects and emotions of convolution. The chapter explores the work that political and social misdirection’s orchestrating disbelief *do* in relation to our capacities to perceive normative structures of deception, and both their refusals and possibilities for new forms of world-making and practices of survival.

The chapter provides a bridge between analyses of the structures of non-physical abuse and the transmutation and convolution, common to specific operations of power, and their affective and emotional dimensions. The chapter explores how transmutational logics operate through reversal, substitution, displacement, and distortion, multiplying realities, with the potential to create discordance and disbelief. However, disbelief as a political emotion also performs parallel work, where the convolution is hidden by appealing to a shared commons. This includes appeals to fictions and fantasies presumed to be held in common, where shocking events, are framed as *interruptions* to normative expectation. Conversely normative expectations are confirmed by appropriating and prefiguring normalized emotional and affective responses as part of disbelief’s communicative

structures. Confabulation is part of the cultural politics of disbelief, related to the imagination of “what if” scenarios, exploited by forms of power that hide their operations through psychologizing and individualizing tendencies.

Chapter 5, *Cover Stories*, explores the histories and dynamics of two genres of storytelling, smear campaigns and charm offensives, which are common to soft power and non-physical abuse. It debates some of the challenges for activists, campaigners and those wishing to change the tone of conversation around abuse. The challenges relate to how to represent more of the epistemological and affective grey areas of abuse; those that confound the discrete act, that appear in strange reversals, in paradoxical feeling-rules, in malevolent actions and amorphous atmospheres, those that are routinely referred to as non-physical or psychological, low-level, lower risk, below the radar, and where the denials, refusals and incomprehensibility become part of the operations of the abuse. The chapter engages in a form of composite storytelling that expresses the workings of the techniques that link non-physical abuse with information warfare in the context of war, colonial power, and battles for the control of narrative, focusing specifically on the origins of the smear campaign and charm offensive.

The tactics are brought to life through the viewpoint of my fictional character Ada. Ada, like many other targets of narcissistic abuse has googled some of the tactics to try to make sense of their experiences. Smear campaign is a common search term, as is pathological lying and the charm offensive. Ada is also a fictional character whose stories are shaped through unreliable narration, a response to the ways in which targets are positioned within perpetrator stories as unstable, mistaken and misguided in their beliefs. The chapter develops the concept of “abuse assemblages” to move beyond abuse as a discrete act, as a way of apprehending the work that different media practices do within *transmutational* economies of abuse. This includes where abuse might appear as a specific matter of concern, often disappearing and reappearing as something else as it moves across different settings and contexts. This includes matters of abuse being recast through marketing logics, reappearing as reputation, brand image, corporate interests, and political capital. The chapter refigures the social figure of the abuser as a *media operator* engaging in forms of soft power aligned to the management of reputational economies.

Chapter 6, *Post-Truth*, opens with a discussion of a novel that has rewritten Kafka’s *The Trial* for the “age of gaslighting” allowing me to foreground some of the issues that the chapter explores in the context of the psychopolitics of post-truth. Developing the concept of *spaceship moments*, I explore the agency of incongruities in the context of confabulations of reality that link political lies with the structures of non-physical abuse. Situated within an overview of some of the different ways in which gaslighting as a structure of communication and power has been developed across a range of interdisciplinary fields of study, I explore the generativity of gaslighting as a form of mediality that is not new but has become part of how media power has adapted to changing media ecologies. It has been argued by many media and cultural theorists that practices of fakery, dissimulation, and other techniques of deceit and deception have significantly changed the media frame (see for example, Carmi, 2020; Juhasz, Langlois and Shah, 2021). Gaslighting and its different iterations, are situated within a conjuncture where there is a recognition that we need analyses of forms of power that operate through decentring the rational subject, and through a communicative

apparatus that is dilemmatic, inhabiting doubt, ambiguity, and forms of governance that undermine epistemic authority.

Exploring the affectivity of grey media, the chapter analyses two “disobedient tweets” that became actors within structures of communication and networks of relations revealing the fugitive and virtual potentials they can carry and attract. The focus on small data helps to zone in on the small acts that become part of networks of affect that are attached to the symbolic energy driving battles for the control of narratives. Statements contained in Tweets can operate as *primes*, exploiting situated action and possibilities, that have the potential to activate a range of different beliefs, opinions, and feelings, contributing to the convolution characteristic of grey forms of mediation. The chapter also explores how primes can also backfire, revealing that there is always the possibility of other narratives that cannot be anticipated or pre-empted in advance. In this sense media events are always more-than-one, revealing the multiplicitous, layered, distributed and generative nature of any attempt to anticipate the future.

Chapter 7, *AI Deception* develops an approach to confabulating and convoluting realities through taking AI Deception as a case study. Questions of what makes us human and therefore different from machines have moved from a focus on emotion, affect, and the nature of intelligence in earlier debates about AI (see Kember, 2003; Wilson, 2010) or expanding cognition to include a “cognitive nonconscious” (see Hayles, (2017), to a focus on forms of human intuition and anticipation and their trained and automated algorithmic forms (Pedwell, 2023). Alongside the development of human-machine intelligences, questions of what it means to be rational and sane, and whether AI is capable of the capacity to dissimulate, err, deviate, or even hallucinate has become a key contemporary question. The moral, ethical, and normative assumptions framing how we understand automated deception, for example, has become one of the key debates shaping what in alarmist terms has been called AI’s threat to extinction or so-called “AI’s Doomsday”ⁱⁱⁱ. Within the context of AI Deception, the normative consequences of what has been popularized as “robots that can lie” or hallucinate have been marked out as requiring much more attention, with cognitive psychologists leading some of the key research directions (see Kneer, 2021).

The author’s approach to some of these questions is aligned to work across the humanities on questions of data justice and algorithmic bias, with a focus on those psychological assumptions that are being made about deception and learning to lie that are being imported into some of the debates about AI deception. The chapter grapples with a number of vexed questions surrounding AI Deception, challenging some of the psychological assumptions that are made about the nature of thinking, conscience, and empathy, that are part of the psychology of false beliefs and ToM (theory of mind). In different ways these areas of psychology are framing how deception is framed and understood within the context of AI Deception and post-truth politics. The psychology of false beliefs is brought into dialogue with critical work across the humanities on data politics, automation, intuition in machine-learning environments, and the Alien logic of AI. The chapter focuses on the malign forms of affective-cognitive sensing that are becoming automated in machine-learning environments, and their relationship to forms of sensing that challenge neuro-cognitive models of mind and rationality. The chapter draws inspiration from a song

released in 1981 by the *Tom, Tom, Club*, “Wordy Rappinghood” to explore some of the issues at stake.

Chapter 8 – A *Biopolitics of Convolution*, concludes the book by outlining a biopolitics of convolution as an ensemble of power-people-politics characteristic of the present conjuncture. The term *biopolitics of convolution* is developed to refer to a biopolitics for governing populations that brings together two forms of power that draw from the psychological and psychiatric sciences, PSYCHOpower (Orr, 2009) and psychiatric power (Rose, 2019). They are both becoming normalized and entangled within systems of power associated with post-truth, and the automation of specific practices of emotional reasoning specifically associated with the more malign aspects of confabulation. The chapter explores and contextualizes the biopolitics of convolution within critical mental health activist challenges to the concept of disorder, a key mechanism of psychiatric power. It draws together some of the themes of the book and explores the implications of the argument for our understandings of sanity, rationality, morality, evil, and what counts as disorder and deviance, particularly as they are constituted and unsettled within liberal humanistic understandings.

Grey Media

Gaslighting Post-truth AI Deception

Lisa M Blackman

Contents:

Introduction: Confabulating and Convoluting Realities

- : Stranger than Fiction
- : Doublespeak
- : The Triumph of Small Moments
- : Grey Media
- : Convolution
- : Experts-by-Experience
- : *Nonsense*
- : Overview and Chapter Breakdown

Chapter 1: Gaslighting-as-Media

Introduction: The Nonrational and Genre flailing

- : Disbelief
- : The Mediality of Gaslighting
- : Confabulation
- : Grey zones of Power
- : Abduction
- : Battles for the Control of Narrative
- : The “Master Manipulator”
- : PSYCHOpower
- : Evil Media
- : Conclusion

Chapter 2: Nonsense

Introduction: Sense and Nonsense

- : The Psychology of Confabulation

- : Knots
- : Dispelling
- : Traps
- : Cognitive “Flexibility”
- : Social Cognition
- : Unfeeling-work
- : Counterfactual Reasoning
- : Speculative Methods
- : Conclusion

Chapter 3: Information Warfare

Introduction: Down the Rabbit Hole

- : *Weak* Ethnographies
- : Unreliable narration: An Affect-laden scene
- : Technologies of Subjugation
- : Epistemic Malevolence
- : Affective Practices
- : Rage-shame
- : *Vignette*
- : DARVO
- : Reactive Abuse
- : “Textbook Narcissism”
- : It’s Fiction like a Sitcom
- : Anticipation
- : Conclusion

Chapter 4: Disbelief as Political Emotion

Introduction: What if?

- : Disbelief
- : Emotions and Affects of Convolution
- : Broken Narratives
- : “Craziness” as methodology
- : Conclusion

Chapter 5: Cover Stories

Introduction: Ways of telling

- : Abuse Assemblages
- : The Charming Man
- : Charm Offensive
- : Media Operatives
- : Affective Atmospheres
- : Smear Campaign
- : *Smearing*
- : Reflexive Control
- : Conclusion

Chapter 6: Post-Truth

Introduction: The PSYCHOpolitics of Post-Truth

- : Changing Media Frames
- : Political Gaslighting
- : Cultural Gaslighting

- : Spaceship moments
- : Jiggery Pokery
- : Paradoxical Affective Couplings
- : Triangulation
- : Gaming and Gusto
- : Conclusion

Chapter 7: AI Deception

Introduction: Learning to Lie

- : What are words worth?
- : Speculative machines
- : Hallucinating AI
- : The “psychology” of false beliefs
- : Robots that Lie
- : Poisonous Pedagogies
- : Mind-reading
- : Artificial Empathy
- : Epistemic Injustice
- : Gut sensing
- : Intuition in machine-learning environments
- : Conclusion

Chapter 8: A Biopolitics of Convolution

Introduction: Pathocracy

- : Robert Mugabe – a hero and a tyrant

- : Evil and Psychopathology
- : A Biopolitics of Convolution
- : Politics of Mental Health
- : Drop the *disorder*: Critiques and Challenges
- : The psychopath as a social figure
- : Media Cultures
- : Beyond Binaries
- : Moral insanity and the Nonrational
- : Conclusion