

Literary Language and Cognitive Experience: How the Reading of English Literary Fiction Shapes Empathy, Theory of Mind, and Social Cognition

Dr. Chander Mohan

Associate Professor, Head of the Department of English at GDC Majalta
chandermohanpant@rediffmail.com

Received: 03/06/2026 | Revised: 08/06/2026 | Accepted: 20/06/2026 | Published: 29/06/2026

Abstract

This paper explores the intersection of literary language, cognition, and social understanding, examining the empirical evidence for the hypothesis that sustained engagement with English literary fiction cultivates empathy, enhances Theory of Mind (ToM) capacities, and reshapes social cognition more broadly. Synthesising findings from cognitive poetics, developmental psychology, and experimental aesthetics, the paper argues that the particular textual features of literary fiction including deep point-of-view narration, affective free indirect discourse, moral ambiguity, and defamiliarisation create distinctive cognitive demands that exercise and refine the mental capacities underlying interpersonal understanding. The paper also addresses methodological objections to this claim, considers the role of genre and text difficulty, and discusses implications for reading pedagogy and the place of literary education in cognitive and social development.

Keywords: *cognitive poetics, empathy, Theory of Mind, literary fiction, English literature, social cognition, defamiliarisation, free indirect discourse*

1. INTRODUCTION

1.1 The Literary Hypothesis

The proposition that reading literature makes us more empathetic more finely attuned to the inner lives of other people has a long history in humanistic thought. Writers and critics from Samuel Johnson to Martha Nussbaum have argued that literary narrative, by inviting readers to inhabit perspectives radically different from their own, cultivates a form of moral imagination essential to ethical life. For much of its history, this argument remained in the realm of assertion: compelling as a philosophical claim, but without empirical grounding. Over the past two decades, however, researchers working at the intersection of cognitive psychology and literary studies have begun to subject the literary hypothesis to experimental scrutiny. The results, while contested and methodologically complex, constitute a body of evidence that demands serious engagement from scholars of English literature, applied linguists, and educators alike. The question is no longer simply whether literature matters, but precisely how the distinctive features of literary language engage and modify the cognitive processes that underlie social understanding.

1.2 Scope of the Present Study

The present paper focuses specifically on English literary fiction, for two related reasons. First, the majority of the relevant empirical research has been conducted with English-language materials and English-speaking participants, making it necessary to attend carefully to the question of whether findings generalise across languages and literary traditions. Second, the particular formal features of the English literary tradition specifically, the development of the psychological novel from Samuel Richardson through Jane Austen to Henry James and beyond provide a distinctive and especially rich site for examining the relationship between literary technique and cognitive engagement.

The paper proceeds as follows. Section 2 reviews the psychological concepts of empathy and Theory of Mind and their relationship to reading. Section 3 examines the formal properties of literary language that are hypothesised to drive cognitive effects. Section 4 surveys the experimental evidence and addresses key methodological objections. Section 5 considers implications for literary pedagogy and concludes with an assessment of the current state of the field.

2. EMPATHY, THEORY OF MIND, AND NARRATIVE

2.1 Definitions and Distinctions

'Empathy' is a term that has been deployed with considerable variability in both popular and scholarly discourse. It is important to distinguish at least three senses in which the word is used in the relevant literature. Affective empathy refers to the capacity to feel what another person is feeling to be moved by another's grief or to share in another's joy. Cognitive empathy refers to the capacity to understand what another person is thinking or feeling, without necessarily sharing that emotional state. Perspective-taking refers more specifically to the deliberate cognitive process of imagining oneself in another's situation.

Theory of Mind (ToM) sometimes referred to as 'mindreading' or 'mentalising' denotes the broader capacity to attribute mental states, including beliefs, intentions, desires, emotions, and knowledge, to other agents, and to use those attributions to explain and predict behaviour. ToM is typically assessed using false-belief tasks: tests that require subjects to recognise that another agent holds a belief that differs from reality and to use that recognition to predict the agent's behaviour. Deficits in ToM are associated with autism spectrum conditions and certain forms of social anxiety, and conversely, high ToM performance is associated with greater prosocial behaviour and interpersonal competence. The claim that literary reading enhances empathy encompasses both affective and cognitive dimensions, but the cognitive dimension the idea that literary fiction specifically trains ToM has attracted the most sustained experimental attention, partly because ToM is more readily operationalised and measured than affective empathy.

2.2 Narrative as a Cognitive Laboratory

Jerome Bruner (1986) famously distinguished two modes of cognitive functioning: the logico-scientific or 'paradigmatic' mode, concerned with abstract truth claims and causal explanation, and the narrative mode, concerned with human experience, intentionality, and the particularities of time, place, and character. Bruner argued that narrative thinking is the primary means by which human beings make sense of social life that we understand persons and their actions through story rather than through logical analysis.

This insight has been developed in subsequent decades by cognitive narratologists, who have argued that narrative comprehension is itself a form of social cognition. To follow a story, readers must track multiple agents with distinct and sometimes opposed motivations, monitor shifting alliances and enmities, maintain models of what each character knows and does not know at any given point in the narrative, and integrate this information with their own prior knowledge of human psychology and social convention. These demands, it is argued, provide exercise for precisely the same mental systems that underlie ToM performance in social life.

"Narrative simulates social experience in a way that allows the reader to rehearse the cognitive and affective skills required for real-world social interaction." — Mar & Oatley (2008, p. 173)

3. FORMAL PROPERTIES OF LITERARY LANGUAGE

3.1 Point of View and Deep Characterisation

The most fundamental formal feature hypothesised to drive ToM effects in literary fiction is the use of deep interiority the sustained representation of characters' inner mental lives. Unlike the discourses of everyday social life, in which we infer others' mental states from external behaviour, literary fiction typically provides direct, elaborated access to characters' thoughts, feelings, desires, doubts, and memories. Readers of Tolstoy's *War and Peace* or Virginia Woolf's *Mrs Dalloway* are immersed in streams of consciousness that provide a richness of psychological detail unavailable in ordinary social interaction. The English novel has been particularly associated with the development of techniques for representing interiority. From Samuel Richardson's epistolary novels, in which characters reveal their inner lives through letters composed in moments of high emotional intensity, through Jane Austen's ironic deployment of free indirect discourse, to the stream-of-consciousness experiments of Joyce, Woolf, and Faulkner, English literary tradition has produced an exceptionally diverse repertoire of narrative devices for mediating between the external surfaces of social life and the complex inner experience of individuals.

3.2 Free Indirect Discourse

Free indirect discourse (FID) the technique by which a narrator's voice blends imperceptibly with a character's inner voice, producing sentences that are simultaneously the narrator's report and the character's thought is perhaps the most cognitively demanding and theoretically significant of these devices. Jane Austen's fiction provides some of the canonical examples in English literature: sentences such as 'It was a truth universally acknowledged that a single man in possession of a good fortune must be in want of a wife' function simultaneously as ironic authorial observation and as a representation of the social doxa internalised by the characters of the novel. Psycholinguistic research has demonstrated that the processing of FID makes particular demands on readers' ToM capacities, because successful comprehension requires the simultaneous maintenance of two distinct perspectives the narrator's and the character's and the tracking of where one ends and the other begins. Zunshine (2006) has argued, drawing on cognitive theory, that literary fiction systematically requires the tracking of multiple 'embedded intentionalities' of this kind, thereby exercising and developing ToM in ways that simpler narrative forms do not.

3.3 Defamiliarisation and Aesthetic Disruption

Victor Shklovsky's concept of defamiliarisation (*ostranenie*) the technique by which literary language disrupts automatic perception, compelling the reader to see familiar objects and situations afresh has been revisited by cognitive poetics as a potential mechanism for the social cognitive effects of literature. The argument is that literary works that disrupt readers' habitual expectations of character, social situation, and narrative event require more extensive and effortful mental modelling than works that confirm genre conventions, and that this greater effort may produce stronger effects on cognitive flexibility and perspective-taking. Empirical work by Kidd and Castano (2013) provided initial experimental support for the idea that literary fiction, as opposed to popular genre fiction or non-fiction, produces distinct effects on ToM performance. Participants randomly assigned to read excerpts from literary fiction scored higher on the 'Reading the Mind in the Eyes' test a standard measure of cognitive empathy than those assigned to read popular genre fiction or non-fiction. The study attracted enormous attention and has been widely replicated, though not without controversy.

4. EXPERIMENTAL EVIDENCE AND METHODOLOGICAL DEBATES

4.1 The Kidd and Castano Studies

Kidd and Castano's original 2013 paper in *Science* reported five experiments in which short-term exposure to excerpts from literary fiction produced improvements on the Reading the Mind in the Eyes Test (RMET) relative to popular fiction, non-fiction, or no-reading control conditions. The paper became one of the most cited articles in the social psychology of reading and generated extensive popular media coverage. It also prompted a substantial replication effort, the results of which have been mixed. Several direct replications found no significant effect of literary versus popular fiction on RMET scores. A large pre-registered multi-site replication (Samur et al., 2018) failed to reproduce the original findings with a substantially larger sample. Subsequent meta-analyses have produced inconsistent conclusions about the overall magnitude and robustness of the effect, reflecting both genuine heterogeneity in study designs and the well-documented difficulties of replicating small-sample social psychology experiments. Kidd and Castano have responded to these critiques by arguing that the null replications employed less prototypically literary texts or less sensitive outcome measures, and by refining their theoretical account to identify more specific predictions. The debate continues, and it illustrates broader tensions in the psychology of reading between laboratory methods suited to causal inference and the ecological validity of real-world literary experience.

4.2 Long-Term and Naturalistic Evidence

The bulk of the experimental evidence for literary fiction effects has focused on short-term, laboratory-based exposure typically, reading a single text excerpt over a brief experimental session. Critics have noted that this design is poorly suited to the hypothesis that sustained, habitual literary reading over years or decades produces cumulative effects on social cognition. The relevant counterfactual is not 'how does reading a page of Chekhov compare to reading a Wikipedia article' but 'how does a lifetime of serious literary engagement compare to a life with little fiction reading?' Evidence bearing on the long-term hypothesis comes primarily from correlational and survey research. Mar et al. (2006, 2009) found that lifetime fiction reading, measured through an author recognition test, was positively correlated with performance on social cognition tasks, even after controlling for verbal intelligence. These correlations are consistent

with a cumulative effects hypothesis but cannot establish causation: it is possible that individuals with high social cognitive abilities are drawn to fiction reading rather than that reading causes high social cognition. Developmental research offers a partial bridge between correlational and experimental evidence. Studies with children and adolescents have found that home literacy environments — the presence of books, frequency of reading aloud, quality of book-mediated conversations about characters' thoughts and feelings — predict ToM development above and beyond general cognitive ability (Adrian et al., 2005; Symons & Clark, 2000). These findings suggest that the relationship between narrative language and ToM may be stronger and more developmentally significant than short-term adult experiments can readily reveal.

5. IMPLICATIONS FOR LITERARY PEDAGOGY

If the literary fiction hypothesis is even partially correct if sustained engagement with the kinds of texts central to the English literary tradition does cultivate social cognitive capacities of genuine importance then the progressive marginalisation of literary reading in curricula shaped by utilitarian metrics deserves serious reconsideration. The reduction of English literature to a set of examination texts processed for plot summary and key quotation identification, divorced from the kind of close, reflective, empathically engaged reading that the hypothesis requires, may undermine the very mechanism through which literature is supposed to be educationally valuable. This does not mean that literature teachers must become cognitive psychologists or that the value of literary education should be reduced to its putative social-cognitive effects. The argument for literary education has always encompassed dimensions of aesthetic experience, cultural memory, ethical reflection, and linguistic development that cannot be captured by a ToM measure. But the cognitive evidence does provide an additional, empirically grounded argument for the importance of sustained literary engagement one that may carry weight in policy contexts where purely humanistic arguments are vulnerable to dismissal as elitist or economically irrelevant. Practically, the evidence suggests that pedagogy which encourages deep engagement with literary characters' interiority attending carefully to point of view, exploring free indirect discourse, debating the plausibility and coherence of fictional minds is more likely to produce cognitive effects than pedagogy focused primarily on plot, theme, or historical context. Approaches drawing on cognitive poetics, such as those developed by Stockwell (2002) and Burke (2011), offer practical frameworks for bringing cognitive awareness into literary classrooms in ways that do not compromise the richness of literary interpretation.

6. CONCLUSION

The question of how literary language shapes cognition is both genuinely important and genuinely difficult. It brings together literary scholarship, cognitive psychology, developmental science, and philosophy of mind in ways that require methodological pluralism and careful attention to the limits of any single research tradition. The evidence reviewed in this paper does not settle the literary hypothesis, but it establishes its credibility as an empirical claim worthy of sustained investigation. What is clear is that literary fiction and the English novel in particular, with its centuries-long development of techniques for representing the depth and complexity of inner life makes cognitive demands on its readers that are qualitatively different from those made by non-fictional or genre-fictional text. Whether those demands translate into durable effects on social cognition depends on variables text selection, reading habits, age of first exposure, quality of engagement that future research must work to specify. For scholars of English literature and language, the most productive response to this research tradition is neither uncritical enthusiasm nor defensive dismissal. It is, rather, the kind of careful, disciplined, imaginative engagement with difficult questions that good literary reading itself models and requires. In attending to what reading does to minds, we may come to understand better what minds bring to reading.

REFERENCES

- Adrian, J. E., Clemente, R. A., Villanueva, L., & Rieffe, C. (2005). Parent-child picture-book reading, mothers' mental state language and children's theory of mind. *Journal of Child Language*, 32(3), 673–686.
- Bruner, J. (1986). *Actual minds, possible worlds*. Harvard University Press.
- Burke, M. (2011). *Literary reading, cognition and emotion: An exploration of the oceanic mind*. Routledge.
- Kidd, D. C., & Castano, E. (2013). Reading literary fiction improves theory of mind. *Science*, 342(6156), 377–380.

Literary Language and Cognitive Experience: How the Reading of English Literary Fiction Shapes Empathy, Theory of Mind, and Social Cognition

Dr. Chander Mohan

- Mar, R. A., & Oatley, K. (2008). The function of fiction is the abstraction and simulation of social experience. *Perspectives on Psychological Science*, 3(3), 173–192.
- Mar, R. A., Oatley, K., Hirsh, J., dela Paz, J., & Peterson, J. B. (2006). Bookworms versus nerds: Exposure to fiction versus non-fiction, divergent associations with social ability, and the simulation of fictional social worlds. *Journal of Research in Personality*, 40(5), 694–712.
- Mar, R. A., Oatley, K., & Peterson, J. B. (2009). Exploring the link between reading fiction and empathy: Ruling out individual differences and examining outcomes. *Communications*, 34(4), 407–428.
- Samur, D., Tops, M., & Koole, S. L. (2018). Does a single session of reading literary fiction prime enhanced mentalising performance? Four replication experiments of Kidd and Castano (2013). *Cognition and Emotion*, 32(1), 130–144.
- Stockwell, P. (2002). *Cognitive poetics: An introduction*. Routledge.
- Symons, D. K., & Clark, S. E. (2000). A longitudinal study of mother-child relationships and theory of mind in the preschool period. *Social Development*, 9(1), 3–23.
- Zunshine, L. (2006). *Why we read fiction: Theory of mind and the novel*. Ohio State University Press.