The Practicality of ToM as a Cognitive Approach in Literary Criticism

Yingjun Deng (Alva)

Department of Social Arts, The University of Sydney, 2050, Australia

alvadeng@qq.com

Abstract

This essay will discuss what condition should be considered and applied while using Theory of Mind (ToM) in critical reading or literary interpretation with a generic discussion on autobiography. I argue that the practicality of ToM in reading literature is not equally functional in different genres, and further suggesting that the application of ToM in literary reading should be deployed accordingly by different genres. Such categorisation will allow us to get a more detailed understanding of ToM's role in literary criticism.

Keywords

Theory of Mind, critical reading, literary studies.

1. Introduction

Theory of Mind (ToM), a term used by cognitive psychologists to indicate an inherent human capacity of mind reading, interpreting the mind of others in the social environment by reading and interpreting verbal or nonverbal representations. Literary critics such as Lisa Zunshine, Alan Palmer, Mark Turner, and Keith Oatley have theorised ToM as a cognitive approach in studying/reading literature from different dimensions. Two major figures in this field, Alan Palmer and Lisa Zunshine have contributed comprehensive discussions on the application of ToM in literature, suggesting its usefulness in studying literature as a whole. Zunshine examines ToM as the major reason for triggering and maintaining readers' interest in literature, considering the complex mental states knitted in novels is a 'chess game' designed by writers to play with readers' ToM. And Palmer considers ToM as a cognitive mechanism to explore the sociocultural value hiding behind the texts, arguing that there is a 'social mind' in both writing and reading fiction. The basic and shared idea of their theories is the hypothesis that both writers and readers attribute fictional minds as the same way they attribute actual minds. Though both Zunshine and Palmer have acknowledged that it does not apply to all area, they have not probed deep into the issue and provided a more detailed account of the limitation and possible solution. Such claims are foregrounding, but too general to be practical for literary critics. Therefore, this essay will explore this issue further to speculate what condition should be considered and applied while using ToM in critical reading or literary interpretation with a generic discussion on autobiography to contextualise the discussion. I argue that the practicality of ToM in reading literature is not equally functional in different genres, and further suggesting that the application of ToM in literary reading should be deployed accordingly by different genres. Such categorisation will allow us to get a more detailed understanding of ToM's role in literary criticism.

2. The Application of Theory of Minds

2.1 Recognising the Factuality and Fictionality

Before cutting into my main argument, I want to discuss a major issue that I found significant in both understanding the value of ToM and the necessity of applying it in critical reading by different genres. The value I am discussing here specifically pointing to the literary value that has potentials to make a contribution to literary criticism, rather than a value merely targeting the reading process of the majority of readers. The discussion will allow me to explain further why I choose to exemplify my argument with the genre of autobiography and its relation to ToM. This point runs opposite to what Zunshine has recognised in her theory that reading fictional minds is highly similar or even equal to

mind-reading in real life. Arguably, the situation related to literature is quite different, and it is important to address the issue in order to make such cognitive approach function properly; that is, fictional minds is not the same with real minds, and the recognition of the difference is the critical point that makes reading fiction evolve from the level of understanding to appreciating (critically reading) literary narratives or artefacts as a whole. In her book, Zunshine has provided substantial evidence in illustrating its indispensable role in 'understanding' fiction, but barely emphasises its significance as a means of literary criticism. Similarly, Palmer has stressed that in order to understand literature, we need to clarify both the similarity and difference between fictional minds and real minds, though literary critics of the field including himself have been putting more attention on its similarities. (Palmer 206) It seems like focusing on their similarities is an inevitable way for cognitive literary critics to take in order to legitimise their approaches in literary studies. Likewise, Mark Turner argues that because narrative schemes mirror human life in so many ways, to read literature as it assimilating real life is to practice mind-reading at a deeper level. (Turner, as cited in Oatley 19) However, to recognise the distinction between fictional minds and real minds is equally important to the practice of ToM in critical reading at a deeper level because for readers, especially literary critic who is familiar with various approaches of literary criticism and basic elements that construct a literary text such as characterization or focalization, they tend to be more sensitive to different components of a text than they do in real life situation. Therefore, the effortlessness and unconsciousness of mind-reading capacity in reading literature that have been highlighted by critics such as Zunshine, Turner, and Palmer are not accurate in such context, though it still holds true to its function as a fundamental mechanism in understanding literature. In other words, to appreciate literature or to perform literary criticism with ToM requires more 'embedded levels of intentionality'[1] because of the differences. In this regard, critical reading of literary works is conducted both consciously and unconsciously. The overemphasis of similarities by major critiques of this field to some extend invalidates ToM's practicality in reading literature, becoming one of the major limitations of such approach.

Why fictional minds cannot be considered as the same as real minds and why it is not valid to theorise fictional minds as real minds without equally emphasising their differences? Firstly, it is because literature as literary artefacts is both reflective and contradictory to real life; by nature, literature is more complicated than real life situation. Here I draw on the notion of literature that is universally accepted as works of art with aesthetic features. From formalist points of view, as Shklovsky holds, the purpose of art is to 'defamliarize' objects in life that have been familiar or taken for granted, and the technique of art is to 'increase the difficulty and length of perception because the process of perception is an aesthetic end in itself and must be prolonged'. (Shklovsky 754) Also, from phenomenological point of view, as Iser holds, in the process of reading, surprise and frustration appear as the betrayals of readers' expectation, generating the enjoyment of reading. (Iser 289-292) Though limited, their theories tell us how important the technique of defamiliarisation/certain degrees of novelty is in literature. The idea of defamiliarisation echoes the way sophisticated writers play with readers' mind-reading capacity, creating more than three levels of 'sociocognitive complexity'. (Zunshine 13-18) In other words, by playing with ToM, writers defamiliarise their works from every day's normality by challenging readers with the presentation of multiple mental states, and such writings with multiple-embedded intentionality can be counted as 'fiction' (as Zunshine theorises the third level as the baseline for fiction). That is also what critics, especially Zunshine and Palmer rely on as a theoretical ground to reason their practices of ToM in literary interpretation, which is credible and successful in theorising ToM as a cognitive approach in literature. However, the word 'Defamiliarisation' also emphasises the strangeness of literature as an art that differentiates it from reality, implying that while accepting the aesthetic experiences brought by literary works, readers are also aware of its abnormal nature bought by the artificial. For example, we tend to have more tolerance to fantasy, imagination, or exaggeration in fictions, while in real life, their correspondence would normally be considered as an indication of mental diseases. Why do readers have double

---

[1] A term used by Zunshine to demonstrate the level of ToM involved in understanding fiction

standards towards fictional events and real-life events? It is because of the fact that cognitively we do not treat them as the same. Keith Oatley and Zunshine have discussed ToM from its evolutionary and historical perspectives, emphasising that ToM is a built-in capacity of human being as a special species among other animals on earth. (Oatley 13-26) It reminds us of ToM's engagement with reality, not the imaginary as a built-in mechanism. Therefore, when we see fictional characters and real minds as two highly similar but also different dialogues, it would be more reasonable to term Theory of Mind in literature as a fictional Theory of Mind; that is, ToM in literature is manipulated by authors' techniques such as defamiliarisation that both prompt and hinder ToM functions in the same way as it does in real life. And such manipulation of ToM in literary works sometimes are detectable by readers because of their awareness of fictionality in the reading process.

The awareness of the differences between fictional minds and real minds leads to the second question that why it is not valid to theorize fictional minds as real minds without equally emphasizing both their similarities and differences, I would borrow an example that Zunshine uses to illustrate 'metarepresentational ability' as a source tracking capacity of human beings to process information by tracing back to the utters such as 'who says/thinks …' [2]According to Zunshine, metarepresentation consists of two parts, one is the source of representation as 'the mind behind the sentiment', and the other is the content of representation. It is noted that the absence of metarepresentational ability will hinder us to track the source of information, resulting in merely being able to entertain the content. (Zunshine 47-49) In order to further explain metarepresentation in fiction, Zunshine sets up three scenes based on real-life situation to exemplify. The first scene is a mundane information about weather (A colleague tells you it is raining outside), the second scene contains a rumor serving as a false information (A colleague tells you someone is morally degenerated), and the third scene reports an obviously fake information: a colleague tells you that it is raining golden coins outside. Zunshine concludes that for the final scene, readers might process the information as simply a joke or view the informer (the source of representation) as 'mentally unstable'. (Zunshine 48-51) It demonstrates how our cognitive system can automatically track source in order to form a metarepresentation and how our memories function episodically (store information for future reference). Her conclusion is cognitively and logically acceptable in the real-life context. Zunshine uses this demonstration to explain further the importance of metarepresentational ability in understanding fiction based on the assumption that reading fictional minds is basically the same with reading real minds. Her demonstration is enough to justify the assumption that how we refer to actual minds in ordinary settings is similar to the way we make sense of fictional ones.

Based on her assumptions, I will divert the emphasis on their differences as I have mentioned above to see whether we can get an alternative result by taking both the similarities and differences into account. One should note that we are pre-described by Zunshine's hypothetical situations, unconsciously taking the scene as a real-life situation because the source of representation (Zunshine or her intention of contextualising the argument) is set; that is, Zunshine has informed her readers about her intention of contextualising her discussion. By directly pointing out the source, she allows no rooms for other source of representation as alternatives. However, if we changed the source tag, the interpretation of the scenes might be different: We inform readers that this scene comes from a fantasy novel, one of the fictional characters tells the protagonist that it is raining golden coins outside. In this context, we are aware of the generic convention of fantasy novels and start analysing the situation with ToM that is what I called in the previous paragraph 'a fictional Theory of Mind', a special mechanism with which we practice our mind-reading ability with the awareness of fictionality in reading literature. Thus the ridicule scene would be logically acceptable in the new fictional setting. As Hutto argues, it is true that fictional narratives are mostly built upon familiar sources from daily life, but not always; therefore, such strong claims that Zunshine and Palmer make should be refined or re-termed to a weaker claim: 'fictional minds are modelled on actual minds, at least sometimes.'

---

[2] Metarepresentation, a term used theory of mind psychologists and philosophers of mind, sometimes described as 'a representation of a representation. See Sperber, in *The MIT Encyclopedia*

(Hutto 277) And as Palmer himself puts it, fictional minds are semiotic constructed, it is both similar to and different from real minds. (Palmer 203-206) Therefore, when the reality is fictionalised and made into art, fictional minds are but not always necessarily read as actual minds for two reasons: one is the artistic nature and aesthetic purpose of literature as an artefact, the other is the readers' awareness of its fictionality and artificial nature.

I have reasoned my claim by changing the source tag of Zunshine's example to demonstrate why the analogue of fictional minds and actual minds is not always valid. But this essay aims not to argue against Zunshine and Palmer's theories, but to recognise both the valid and invalid parts and to develop and refine it in one of the possible ways. There are reasons for putting more focus on the similarities; that is, drawing on such analogue legitimises the practice of ToM in reading literature. And such focus does not contradict with their omission on differences, but it does bring limitation to the method because both facticity and fictionality matter to ToM.

2.2 Better Performance: Applying ToM by Different Genres

Zunshine and Palmer have contributes numbers of close readings on literature to theorize their cognitive approaches. But it is observable that both of them have preference to certain types of texts. For Zunshine, she prefers to drawing on works with strong characterisation or unreliable narrators such as Virginia Woolf's *Mrs. Dalloway* (1925) and Nabokov's *Lolita* (1955) and detective novels. For Palmer, he has conducted textual analysis on George Eliot's *Middlemarch* (1874) and Charles Dickens's *Little Dorrit* (1857) in his *Social Minds in Novel*. The selectiveness of their studies implies that their practice of ToM in literature is limited and conditional; it is not applicable to all genres, and its practicality is not equally effective in different genres. As Orley Marron points out, we use ToM in situation that involves human, but it is not adequate to interpret the cognitive process of non-human mind. (Marron 189-90) Similarly, Keith Oatley points out that this cognitive approach shows more useful in some genres while less important in others; for example, it is more applicable in plays, novels and stories than in epic poetry or inward thought in lyric poetry. (Oatley 16-17) Addressing to the issue, I suggest that the practice of ToM in critical reading should be applied accordingly by different genres to function in a more proper and effective way.

But what kind of genres are more suitable for the application of ToM? Based on a few major theories in this field, Zunshine's and Palmer's in particular, this essay tries to list three. Firstly, as Zunshine and Palmer have demonstrated on their studies, it should be the work of fictionality and imagination based on which writers are able to create multiple levels of embedded ToM , setting up a 'chess game' with readers. Secondly, the emphasis should be put on 'sociality', a key concept in Palmer's theory that views the application of ToM in literature from sociocultural perspective that is similar to folk psychology. From this point of view, as Palmer holds, individual identity is built on the interaction with others; therefore both internal and external aspects 'are important components of fictional narratives … just as in real life, where much our thinking is done in groups. (Palmer 211-214) For this reason, 'sociality' in texts is mostly created through the literary technique of characterisation, drawing connection between characters for identity formation. Last but not least, it should be the literary works with comparatively strong source tagging (metarepresentation), in which the information that connects the plot and characters are trackable to readers. To summarize, applying Theory of Mind in literature tends to involve fictional elements, complex theory of mind, characters (sociality), and source tag available to readers to form their metarepresentation. My proposal is not a denial of ToM's function in some genres, but a speculation about in what kinds of genre ToM can be better performed as a cognitive approach of literary criticism. The speculation is based on the assumption that fictional minds are not the same as actual minds, they only assimilate each other in some case, and I especially point to genre in this case. For literary texts that have features mentioned above, the use of ToM as a critical approach on such texts are more applicable and important, therefore more likely to engage with the theory effectively.

2.3 Model Genre: Combining the Factual and the Fictional

To contextualise my argument, I will discuss the generic feature of autobiographical writings, examining it as one of the genres that is more suitable for the practice of ToM. Zunshine and Palmer have not touched upon autobiography in their major studies yet, but they both agree that ToM is not useful in some writing. Hypothetically, this genre has lots of potentials to be a rich ground for practicing ToM.

First of all, the generic features of autobiographical writing is more applicable for ToM. Autobiography nowadays is termed as a hybrid genre, drawing on life while allowing its combination with fictional and imaginative elements. This genre has been testing the boundaries between the factual and the fictional since it was coined in the late 18th century. [3] Its half fiction and half fact feature creates a game for readers to explore between fictional minds and real minds. For example, a typical question from a reader of autobiography would be: whether the narrator/author is lying? Is she/he telling the truth? Brockmeier and Carbaugh suggest that its ambiguity and vagueness about factuality and fictionality creates 'a dynamic that keeps in view actual stories about real life with possible stories about potential life'. (Brockmeier, Jens, and Donal A. Carbaugh 6-9) In other words, writing real person as a character and real life as a story renders a dual nature to the 'I' used in autobiographical writing. When theorising autobiography, Lejeune particularly points out the difficulty of 'I' for readers to perceive; he states that when writer asks 'who am I', it is natural for readers of autobiography to ask: who is 'I', and who it is who says 'who am I'. (Lejeune 8-9) Its generic convention provides readers with complex states of mind to explore whether the 'I' is a reference or an utterance. Accordingly, Zunshine points out that readers 'conjure up an extra mental presence' when we intuitively sense that the narrative is fooling us by disturbing our 'source-monitoring capacities'. (Zunshine 77-79) Therefore, autobiographical narrative in first-person allows readers to practice complex theory of mind for 'I' is a real mind as well as a fictional mind. Additionally, psychoanalysis is frequently applied in the interpretation of autobiographical writings. In their study on Theory of Mind, Premack and Woodruff have noted that creative writers such as novelists have more exotic mental states. (Premack, David, and Guy Woodruff 515) Similarly, Freud also admits that 'the poets and philosophers before me discovered the unconscious. What I discovered was the scientific method by which the unconscious can be studied.' (Freud as quoted in Alvarez 13-16) Therefore, the fact that psychoanalysis is frequently used as one of the major methodologies to interpret life narratives (breaking down multiple states of mind), and the close relation of psychology and autobiographical writing to some extent, prove that the generic nature of the first-person autobiography is a typical representation of consciousness and mind.

Secondly, the writing process of autobiography is another proof of his suitability for ToM. Writers of autobiography aim to craft a self-representation; it often involves with the construction/reconstruction of identity and memories. From Palmer's sociocultural perspective, the construction of identities is realized by two attributions: the internal and the external. (Palmer 210-213) In this genre, the internal points to the subjectivity of the self, and the external points to the subjectivity of others. In other words, autobiographers deploy characterization to reveal how the protagonist perceive himself/herself and how other characters might perceive the protagonist. By doing so, the identity is constructed. Here, characters can be created by imagination to suit whatever identity the author wishes to construct. Thus the interpretation of the text becomes challenging to readers' ToM with the interaction of fictional minds and actual minds for writers can become an unreliable narrator of his/her own stories. Reader's ToM therefore has more material to work on, as Zunshine states, this cognitive capacity of us starts working from suspecting that the narrative is monkeying us around to 'ascribing a whole host of other states of mind ' to that tricky narrative. (Zunshine 78)

Thirdly, the reading practice of autobiography's readers, to some extent, is a typical practice of ToM. Zunshine has pointed out that a quest for truth is a constant motivation that keeps readers going, it is

---

[3] The Oxford Companion to English Literature (7 ed.), autobiography.

a cognitive effort of deciding and cognitive desire of knowing. (Zunshine 70-73) This hunger for truth is more obvious on readers of autobiography. When they pick up a book from autobiography shelf in the bookstore, they are pre-described by their knowledge on generic convention of this genre. Therefore, the reading practice sets out from the expectation of looking for a truth in life narratives in which the fictional are mixed with the true. While autobiographers select different parts of memories and thread them together, the make-up of the text becomes a representation of episodic memories in which the past, like memories, is retrievable. As Palmer puts it, such textual structure enables a creative reading process in which 'reader's strategy is to join up the dots' presented by autobiographers. (Palmer 203-206)

The constant hunt for truth is a constant experiment, both for writers and readers of this genre. Equally important, autobiography provides readers with a source tag that is comparatively easier for tracking. From poststructuralist point of view, instead of taking author as the main agency of the text, reader should be in charge of the interpretation. (Barthes 142-148) Indeed, if we consider author as the source of the text, such practice would oversimplify the situation. Correspondingly, Zunshine holds a positive attitudes towards the poststructuralist view, suggesting that there are alternatives functioning as the source of representation. But she also emphasizes that even though poststructuralist's ideas of 'the death of the author' is possible in terms of source monitoring, it is really not that cognitively feasible because 'there is always an author behind the text' (Zunshine 66-72) Therefore, when the author, narrator, and protagonist are identical in autobiographical writing, readers are provided with a strong source tagging to the author and the genre, making it more convenient for interpretation. Philippe Lejeune terms the practice of autobiographers using their own name as the narrator and protagonist name as 'authorial signature'. (Lejeune 12-14) This idea echoes the concept of metarepresentational tagging in Zunshine's theory. Genre and its generic features as a cognitive metarepresentational tag can change readers' mind in the process of reading and deciding.

3. Conclusion

To conclude, in this essay I have discussed the significance of recognising both the similarities and differences to actual minds and fictional minds and further evaluated the application of ToM in literature by genre, suggesting one of the possibilities with the example of autobiographical writing to contextualise my argument. The main limitation of this essay is it only explores one of the aspects. And correspondently, there is a particular kinds of genres that is not suitable for the practice of ToM need attention. Also, based on my suggestion of practising ToM by different genres, this cognitive approach in literature can be further refined. That is, while it is clear what kinds literary text can or not suitable be interpreted with ToM, genre can be categorised by hierarchy to maximise the literary value of ToM so that this approach can be deployed more effectively. Both Palmer and Zunshine have brought thought-provoking significance for this school of literary criticism, but their approaches might be too general and opened, facing the crisis of subjectivity. Be that as it may, by theorising the subjectivity, theorists of this field have excavated a new ground with lots of potentials, especially for the study of world literature for ToM as a cognitive mechanism is a shared capacity of human beings across culture.

References

[1] Barthes, Roland. "The death of the author (1977)." Image, music, text, essays selected and translated by Stephen Heath (1988): 142-48.

[2] Brockmeier, Jens, and Donal A. Carbaugh, eds. *Narrative and identity: Studies in autobiography, self and culture*. Vol. 1. John Benjamins Publishing, 2001.

[3] Freud, as quoted in Alvarez, Alfred. *The Writer's Voice*. WW Norton & Company, 2005.

[4] Hutto, Daniel D. "Understanding Fictional Minds without Theory of Mind." *Style*, vol. 45, no. 2, 2011.

[5] Iser, Wolfgang. "The Reading Process: A Phenomenological Approach.", 2011.

[6] Marron, Orley K. "Alternative Theory of Mind for Artificial Brains: A Logical Approach to Interpreting Alien Minds." 201.

[7] Lejeune, Philippe. *Le Pacte Autobiographique.* Seuil, Paris, 1975.

[8] Palmer, Alan. "Social Minds in Fiction and Criticism." *Style*, vol. 45, no. 2, 2011.

[9] Premack, David, and Guy Woodruff. "Does the chimpanzee have a theory of mind?." *Behavioral and brain sciences* 1.4, 1978.

[10] Oatley, Keith. "Theory of mind and theory of minds in literature." *Theory of Mind and Literature* (2011): 13-26.

[11] Shklovskiĭ, V. (1992). Art as technique. In H. Adams (Ed.), *Critical theory since Plato* (pp. 750–759). Fort Worth: Harcourt Brace Jovanovich College Publishers, 1992.

[12] Zunshine, Lisa. "Sociocognitive Complexity." Novel: A Forum on Fiction, vol. 45, no. 1, 2012.

[13] Zunshine, Lisa. *Why We Read Fiction: Theory of Mind and the Novel.* Ohio State University Press, Columbus, 2006.