

#### Citation

Andrew piper (2023). "What do characters do? The embodied agency of fictional characters". In: *Journal of Computational Literary Studies* 2 (1). [10.48694/jcls.3589](https://doi.org/10.48694/jcls.3589)

**Date published** 2023-03-20

**Date accepted** 2023-02-24

**Date received** 2023-02-17

#### Keywords

literary characters, fiction, narratology, nlp, theory of mind, embodied cognition

#### License

CC BY 4.0 

#### Reviewers

Fotis Jannidis, Peer Trilcke

#### Note

This paper has been submitted to the conference track of JCLS. It has been peer reviewed and accepted for presentation and discussion at the 2nd Annual Conference of Computational Literary Studies at Würzburg University in June 2023.

#### Article

## What do characters do?

### The embodied agency of fictional characters

Andrew Piper<sup>1</sup> 

1. Languages Literatures and Cultures, McGill University, Montreal, Canada.

**Abstract.** This paper uses machine learning to provide the first ever large-scale estimates of the distribution of actions of literary characters. Learning more about these distributions can help us better understand how different genres, cultures and time periods simulate personhood, i.e. what it means to be an agent in the world. Prior research has emphasized fiction's capacity to promote "social cognition" (Theory of Mind) on the part of readers, where the complex cognitive states of literary characters are thought to facilitate deeper reasoning about human motivations and behavior. The data and models used here add an important dimension to this theory by highlighting how the actions that have increasingly distinguished fictional characters from their non-fictional counterparts over time entail forms of embodiment rather than explicit invocations of cognitive or emotional behavior. These results suggest that embodiment has emerged as a central value in the fictional representation of personhood.

## 1. Introduction

When contemporary writers tell stories, what do their characters *do*? What does the distribution of actions across characters look like and how has this changed from the past? And what can this knowledge about the behavior of fictional characters tell us about the meaning and function of fictional storytelling?

Understanding the actions of fictional characters is important because it can give us access to how personhood – what it means to be an agent in the world – is simulated across time, cultures, and genres. This process of fictional characterization provides insights into the values associated with human or non-human agency (Eder et al. 2010; Frow 2014; Jannidis 2004; Phelan 1989).

Traditions in both empirical and theoretical research have strongly focused on the concept of social cognition when it comes to the role of fictional characters. As Kidd et al. (2016) have written, "We propose that reading fiction can be an exercise in advanced ToM [Theory of Mind] that prompts readers to represent and engage with characters' nuanced mental states." As Palmer (2004) has stated even more emphatically, "narrative fiction is, in essence, the presentation of fictional mental functioning" (5). Similarly, as Anderson has written about the genre of the novel, "The novel has a special capacity and license to convey the phenomenology of the thinking life, and it has demonstrated a special interest in forms of thinking since its inception" (Anderson et al. 2019, 131). The fictional simulation of mental worlds, so this line of thinking suggests, provides

the opportunity to develop more sophisticated forms of social cognition on the part of readers (i.e., Theory of Mind). As Zunshine writes, “We like reading fiction because it lets us try on different mental states and seems to provide intimate access to the thoughts, intentions, and feelings of other people in our social environment” (Zunshine 2006, 25). According to these theories, characters stand at the centre of fiction and minds at the centre of character.

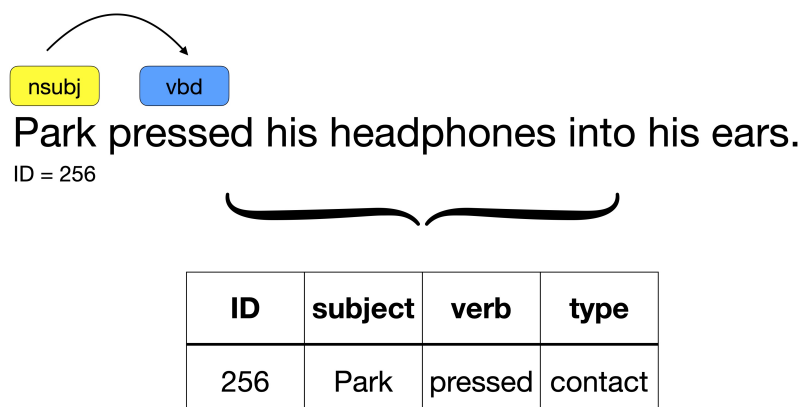
A core challenge for this theoretical framework is the incorporation of knowledge about how such mentalizing actually takes place within fiction. If novels have a special capacity to convey the phenomenology of the thinking life, how is this manifested within the language of novels more broadly? Computational models of text analysis can be useful here to provide more detailed information about the linguistic representation of characters’ actions, and by extension the mental life of characters. This work thus represents a continuation of prior work aimed at understanding the large-scale representation of literary characterization (Bamman et al. 2014; Cheng 2020; Heuser and Le-Khac 2012; Piper 2018; Underwood 2019), with a particular focus on character agency.

In order to estimate the distribution of character actions, this paper utilizes the annotations provided by BookNLP (Bamman 2021). BookNLP is a natural language processing pipeline that includes part-of-speech tagging, dependency parsing, entity recognition, character name clustering, and word super-sense tagging. BookNLP is a particularly valuable resource for this task because it has been trained on literary data (Bamman et al. 2019). In this paper, I use it to identify actions associated with characters and then classify those actions according to higher-level categories based on the super-sense classifications (see Table 1 for a full list). This workflow is then applied to two datasets: the CONLIT dataset, which consists of a collection of 2,754 works of English prose published since 2001 drawn from twelve different genres (Piper 2022), and the Hathi1M dataset (Bagga and Piper 2022), which consists of a collection of 1,671,370 randomly sampled pages of English prose published between 1800 and 2000.

As this paper will show, the actions that distinguish fictional characters from their non-fictional counterparts largely encompass forms of embodiment, such as touching, smiling, shrugging, moving, sensing, etc., rather than explicitly emphasizing cognitive or emotional actions, such as thinking, wondering, or reflecting. Fictional agency distinguishes itself as embodied agency, a fact that has only grown stronger over time. This is not to suggest that fiction is not invested in the representation of inner mental states any more than non-fictional narratives are. But it does suggest that there is a sensori-motor preference surrounding fictional agency that future work on character and social cognition will want to consider more fully. What value does the translation of virtual agency through the human body have for readers?

## 2. Data and Methods

All texts in the two primary datasets mentioned above are first processed through the large model of BookNLP (Bamman 2021). As illustrated in Figure 1, this workflow consists of the following steps for every book:



**Figure 1:** The process of character behavior identification. First the character is identified as either a named-entity or co-referencing pronoun, given a meta-ID and then its grammatical position is identified. If the character is in the subject position (nsubj) then the verb token(s) associated with that character are identified and stored along with the associated super-sense tag.

1. Identifying all “persons” (PER) according to the entity recognition classification system;
2. Assigning an ID to each person using co-reference and character name clustering;
3. Identifying the grammatical position of each person using part-of-speech tagging and only keeping those that are the subject of a sentence;
4. Identifying all main verbs associated with these subjects using dependency parsing
5. Extracting the “super-sense” of these verbs.

Importantly, we use the BookNLP “book” file that allows us to extract the primary actions associated with every character in the subject position and then identify the actions’ type using the associated “super-sense” file. Super-sense tags in BookNLP are generated using a predictive model trained on SemCor, which is based on the taxonomies provided by WordNet’s hypernym trees. Instead of relying on individual keywords for analysis, the super-sense tagging aggregates individual words into more general behavioral categories, but does so using predictive models rather than dictionaries to account for the problem of polysemy. The two sample sentences below illustrate how the model accurately classifies the verb “found” according to different senses, where Example 1 represents a cognitive event while Example 2 represents a perceptual event according to the super-sense taxonomy.

1. “I **found** the work in the small outpatient clinic difficult, as I was certain that many things were getting lost in translation.” [Cognition]
2. “He then **found** himself in a group around a television journalist who had just published his memoirs.” [Perception]

According to the BookNLP documentation, the overall accuracy of super-sense tagging is estimated to be 76%. In the results section, we describe a more fine-grained manual validation exercise, which suggests even higher-level accuracy for the key categories of

| Type          | Top tokens                      |
|---------------|---------------------------------|
| body          | smile, laugh, wear, sleep, feel |
| emotion       | want, like, feel, love, hope    |
| change        | start, begin, get, make, die    |
| motion        | go, come, walk, turn, leave     |
| cognition     | know, think, remember           |
| perception    | see, look, hear, have, feel     |
| communication | say, ask, tell, call, mean      |
| possession    | have, get, find, give, lose     |
| competition   | fight, play, shoot, win, fire   |
| social        | do, try, make, let, work        |
| consumption   | need, use, eat, have, drink     |
| stative       | be, keep, wait, live            |
| contact       | stand, sit, put, pull, open     |
| weather       | light, burn, blow               |
| creation      | make, do, imagine, write        |

**Table 1:** Top tokens for contemporary fiction for each verb type according to the BookNLP super-sense tags.

interest here. [Table 1](#) provides a list of all verb types in the BookNLP model with their most-frequently associated words derived from fiction books in the CONLIT data.

### 3. Results

[Table 2](#) provides an overview of the distribution of fictional character actions in the CONLIT data according to the super-sense schema described above. As we can see in the left column, the most frequent actions undertaken by characters are acts of *communication* followed by *motion* and *cognition*. This gives us a sense of the most common actions undertaken by fictional characters.

However, when we look at the relative frequencies of these actions between fictional and non-fictional narratives measured using a  $G^2$  log-likelihood ratio statistic (Dunning 1993) (right column), that is, when we look at what fictional characters do *differently*, the strongest positive predictors of fictional character behavior are all embodied forms of action (contact, body, perception, motion). Communication, far from being dominant in fiction despite its overall frequency, is actually weakly indicative of non-fiction relative to fictional discourse. Indeed, the only verb types that are statistically distinctive of fictional discourse are the ones that indicate embodied actions (though as we will see embodiment is not limited to just these types). Even if we were to rely on the overall frequencies, the combined frequency of the embodied character actions (contact, body, perception, motion) are almost two times more frequent than cognitive and emotion types combined.

[Table 2](#) gives us some indication of the extent to which the distinguishing qualities of character actions in fiction revolve around behavior associated with different forms of embodiment. Fictional characters spend considerably more time standing, sitting, turning, walking, and smiling than their non-fictional counterparts. On the other hand, they appear to engage in relatively similar levels of explicit mentalizing (knowing, thinking, wanting, hoping, liking).

| Frequency     |           | Log-Likelihood Ratio |                |
|---------------|-----------|----------------------|----------------|
| Type          | Count     | Type                 | G <sup>2</sup> |
| communication | 3,344,071 | contact              | 116,743        |
| motion        | 2,056,877 | body                 | 47,340         |
| cognition     | 1,965,067 | perception           | 39,404         |
| contact       | 1,443,898 | motion               | 33,655         |
| perception    | 1,363,000 | weather              | 67             |
| social        | 945,900   | consumption          | -73            |
| possession    | 896,984   | emotion              | -273           |
| emotion       | 796,297   | cognition            | -476           |
| stative       | 717,342   | stative              | -1,457         |
| change        | 619,910   | communication        | -5,226         |
| body          | 599,845   | possession           | -21,834        |
| consumption   | 263,173   | change               | -26,977        |
| creation      | 223,784   | competition          | -30,030        |
| competition   | 89,309    | creation             | -46,578        |
| weather       | 1,754     | social               | -69,318        |

**Table 2:** Overall counts of actions undertaken by fictional characters in the CONLIT data (left column) along with the relative frequency as measured using Dunning's log-likelihood ratio (right column) comparing fictional and non-fictional characters. Positive/negative values indicate actions positively/negatively associated with fiction.

In order to better understand these relationships more broadly across our data, we can aggregate our verb types into two larger classes, one for “embodied” actions and one for “cognitive” and then calculate the fraction of all actions comprised by these types. To do so, I combine the types for *motion+contact+body* for the former and *cognition+emotion* for the latter. I thus frame “embodiment” for the purposes of this paper as a form of corporeal movement and “cognition” as the combination of thinking and emotional feeling. I return to this issue in the discussion section to review limitations and possible alternatives to this approach. The three forms of movement captured here are by no means exhaustive of embodied agency, but they can give us some insights into the nature of the distributions of these kinds of actions across time and genres.

To test the validity of these categories, we manually annotated 500 tokens randomly drawn from the CONLIT data. Tokens were presented in the context of two sentence passages, with the token belonging to a verb in the second sentence. A set of three student readers were then asked: “With respect to the highlighted token is a character a) “physically moving” with any body part, b) “thinking about something or feeling an emotion,” or c) none of the above. A true positive occurs if the super-sense types *motion*, *body*, or *contact* are predicted for a) and the types *cognition* or *emotion* for b). Importantly, for the purposes of this exercise we consider these as mutually exclusive, a point to which I will return in the discussion section. We might think of this as a means of identifying a “primary” understanding of the action for which there may be secondary features (e.g. a movement that indicates a mental state). The validation thus captures the extent to which the super-sense categories align with reader judgments about the behavior of characters.

In terms of our annotators’ agreement, we report a Fleiss’ *kappa* of 0.813, suggesting very high levels of agreement. We note that one reviewer exhibited lower levels of agreement than the other two, with a pairwise *kappa* ranging from 0.75 to .91. In cases where we

did not have total agreement, a fourth expert rater (the author) made the final decision about the categorization of the token (all data is available for review in the provided link). We found that just over one-third of minority votes were ultimately deemed the correct label, suggesting that a simple “majority rule” would be unreliable for testing our model’s accuracy. After reviewing the model’s errors we found five cases where all annotators agreed, but the expert disagreed.

With respect to our model’s predictions, as we can see in Table 3 both categories exhibit reasonably high precision, with slightly more accuracy surrounding the prediction of cognitive states. The lower recall for both categories suggests that actual behavior, whether cognitive or embodied, can be striated across several other super-sense types. Nevertheless, the relative similarity of precision and recall between classes indicates that one class is not significantly more error-prone than the other, suggesting we do not need to take steps to adjust our estimates based on the observed levels of error for the purposes of comparison.

| Type       | Precision | Recall | F1   |
|------------|-----------|--------|------|
| Cognition  | 0.95      | 0.82   | 0.88 |
| Embodiment | 0.92      | 0.80   | 0.86 |

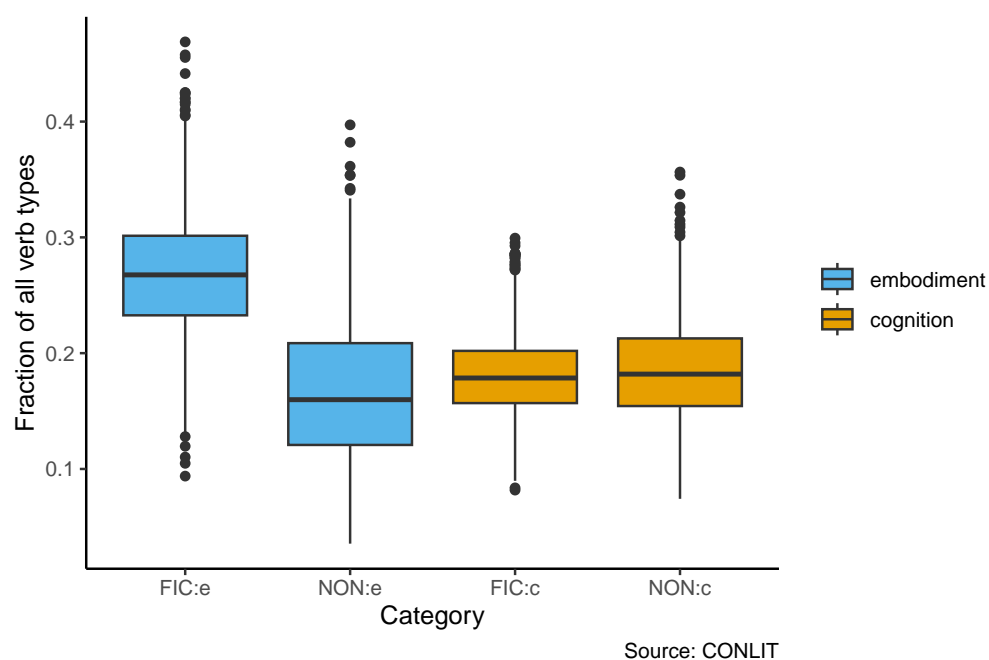
**Table 3:** Accuracy of BookNLP super-sense annotation for the two meta-categories created here: “cognition” (which combines the super-senses *cognition+emotion*) and “embodiment” (which combines the super-senses *motion+body+contact*).

Table 4 shows examples of the way BookNLP classifies verbs under other super-senses than those we designated for each class (i.e. false negatives according to our model). We note that false negatives cluster around *communication* and *perception* for “cognition” and *change* for “motion,” though eight or more total classes are represented in each group’s false negatives suggesting errors are also widely distributed across different types. The verb “see” in particular posed an interesting challenge as it is often used in a cognitive sense, such as “I **see** my death coming” or “In my struggle for objectivity, I **see** myself again.” In these cases, the seeing is an internal mental process, not a physical perceptual one.

| Sentence   | BookNLP Label | True Label |
|--|---------------|------------|
| “He picked it up and <b>took</b> a sip.”                     | consumption   | motion     |
| “I <b>burst</b> into the hall.”                              | change        | motion     |
| “I <b>handed</b> him a cheque.”                              | possession    | motion     |
| “I just can’t <b>find</b> it in me to care that I’m losing.” | perception    | cognition  |
| “You <b>given</b> any thought to what price you’d pay?”      | communication | cognition  |
| “Eric still <b>felt</b> like her secret lover.”              | body          | cognition  |

**Table 4:** Examples of mislabeled actions according to our model. Word in bold is the annotated token.

If we combine our designated super-sense types into our two larger categories, Figure 2 illustrates the disparate impacts of fiction on the prevalence of embodied actions versus cognitive actions in the CONLIT data. Calculating Cohen’s  $d$  on this data shows that the effect size of fiction on the embodied class is substantial (Cohen’s  $d=1.82$ ), while the influence of fiction on the cognitive class is minimal, evidenced by a negligible effect size (Cohen’s  $d=-0.17$ ).



**Figure 2:** Frequency of embodied and cognitive actions in fiction and non-fiction.

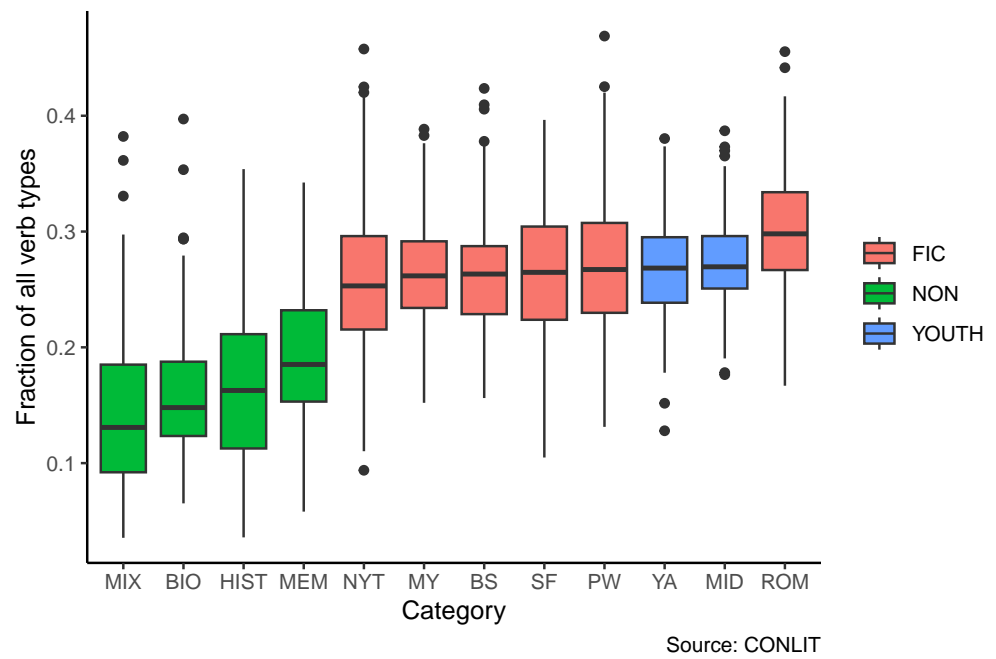
[Figure 3](#) illustrates the inter-genre differences for the embodied class of verbs, highlighting the relative consistency across fictional genres, with the exception of Romance. We note that books written in the first person tend to have slightly higher rates of both embodiment and cognition than those written largely in the third person by about 1 more occurrence every two pages for each type, even when controlling for genre.

Using the Hathi1M data, we can also observe these actions' behavior across historical time. As we can see in [Figure 4](#), actions associated with our embodied verb types have risen considerably over time within fictional prose, while those associated with cognitive and emotional actions have remained largely stable. It is important to note that unlike the CONLIT data, non-fiction in Hathi1M is not exclusively narrative non-fiction but consists of considerable amounts of non-narrative non-fiction, which explains why the cognitive class is considerably lower in this data by comparison with the contemporary data.

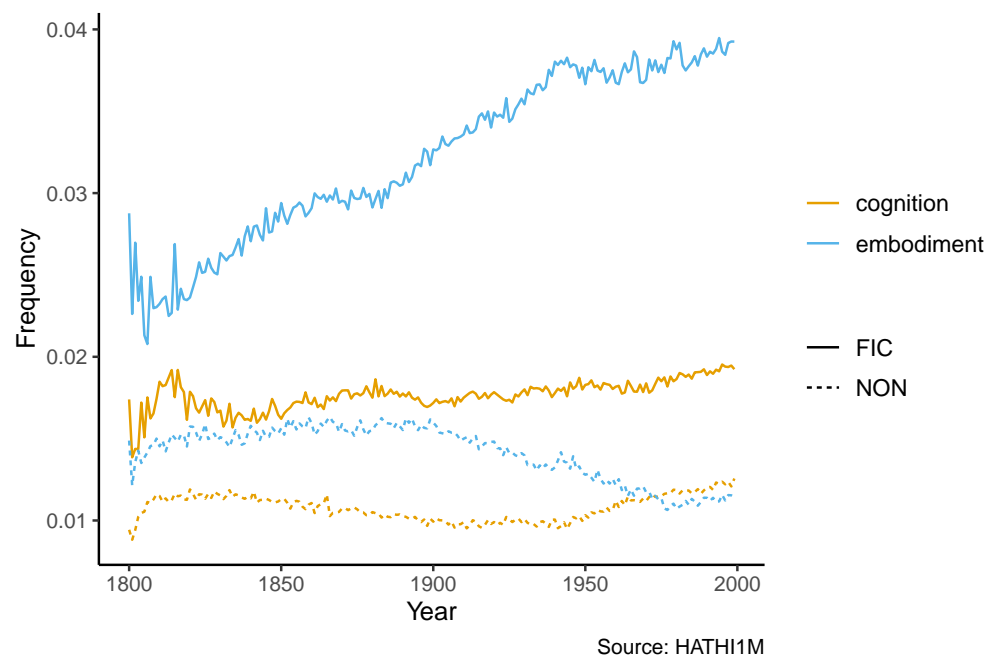
If we break down our historical data back into the individual types ([Figure 5](#)), we see how verbs of *motion* have experienced the largest overall raw increase, but that the relative change across all three classes between the beginning of the nineteenth century and the end of the twentieth is similar across classes (ranging from a 46% increase for motion verbs to a 58% increase for body verbs).

## 4. Discussion

The data and models presented here indicate that the primary way in which fictional narratives distinguish themselves is through a set of embodied actions associated with characters. The distinctive quality of personal agency in fiction is manifested through the agent's body. This is consistent across genres and is also illustrated in the historical data as we see the strongest growth among actions belonging to embodied types of

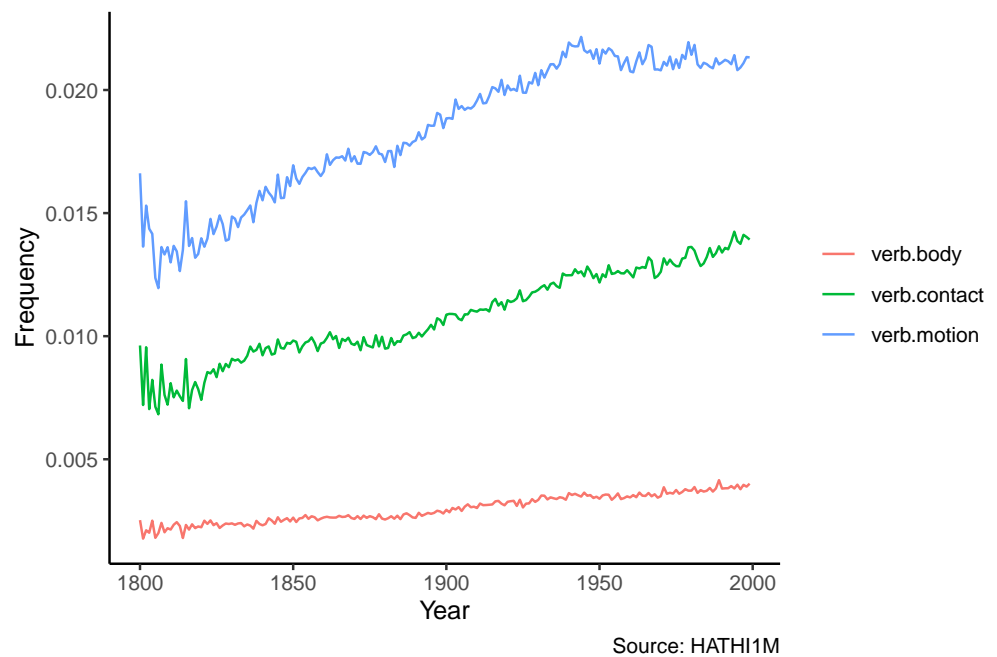


**Figure 3:** Rates of embodied actions across all genres in the CONLIT data.



**Figure 4:** Rate of combined embodiment and cognition verbs for fiction (solid lines) and non-fiction (dotted lines) for the past two centuries.





**Figure 5:** Frequency of the three embodied types in fiction over the past two centuries.

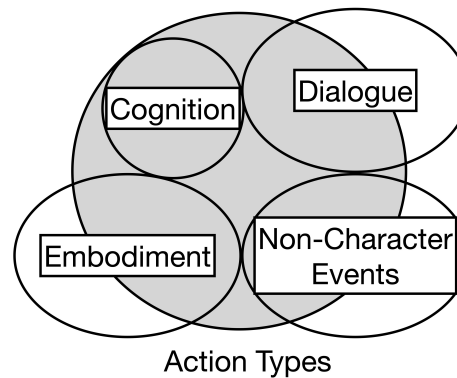
action. It appears that fictional narration, at least in an English-speaking context, has settled on a particular framework of representing agency rooted in the body.

These findings help corroborate prior work on the rising interest in embodiment through the practice of characterization (Heuser and Le-Khac 2012; Underwood 2019). They also raise a host of questions and challenges for future research. One of the main challenges posed by this research is the alignment between super-sense tags and the concepts of “embodiment” and “cognition.” As we saw with the validation exercise above, both embodied and cognitive behavior is captured across a variety of verb types, with meaningful overlap among certain types such as “change” (motion) and “perception” (cognition). While we can be confident given our high precision that we are adequately representing these concepts, we are also missing some of their more widespread uses. Nevertheless, given the very large effect sizes surrounding the differential use of embodiment in fiction and non-fiction, we can be confident about the distinctiveness of fiction’s investment in embodied agency.

On the other hand, the “embodied” actions that are captured by the super-sense tags used here (motion, contact, body) are not necessarily opposed to the act of “cognition” or mentalizing more generally. To return to the opening theoretical framework of “Theory of Mind,” it is safe to assume that when characters are “smiling” or “bursting” into rooms they are conveying something about their internal cognitive states for readers to interpret. As Palmer (2004, 120) argues, “Mental and physical sides of action and behaviour coexist and interpenetrate to the point where they are difficult to disentangle”.

Indeed, potentially *all* actions may contribute to a reader’s mentalizing about a character’s mental state, but with differing degrees of intensity. These could include events not directly undertaken by characters that they directly experience (like floods or fires), a range of action types (motion, possession, etc.) as well as dialogue. Character speech is a key means of revealing their inner states to readers.

## Character Mentalizing



**Figure 6:** Schema of the relationship between event types and character mental states in fictional narratives.

Future models that wish to account for the “phenomenology of the thinking life” within fictional narratives will therefore require new annotation schemes, ones that could potentially incorporate the super-sense tagging scheme used here as well as further relevant features (see Figure 6 for a proposed framework). Such annotation schemes could approach texts at the passage level and assess the intensity or depth of mentalizing represented as well as the accuracy with which readers can “guess” a character’s mental state (agitated, anxious, angry, reflective, etc.). The verb types along with other linguistic features that best predict the assessment of mental depth or state would then help us better understand the features that trigger social cognition, i.e. the ability to mentalize about other people’s mental states that is assumed to be a hallmark of fictional narrative. It would also help draw attention to those literary spaces that are invested in other kinds of representational work. Not all aspects of a work of fiction are about facilitating our thinking about other minds, but rather trigger a range of cognitive states worth exploring.

Overall, the data and models used here provide further evidence that there has emerged a larger consensus around the process of characterization that foregrounds embodied agency at the heart of simulating fictional persons. Whereas prior work has focused on the prevalence of nominal body parts, this work allows us to gain insights around the behavior of characters and its relationship to embodiment. Far from portraying characters as “thinking black boxes,” which readers learn to decode, it seems more appropriate to see fictional narrative in its contemporary form as a cultural technique of modeling “embodied cognition,” i.e. what it means for an agent to be embedded in an environment (Caracciolo and Kukkonen 2021). As Thelen et al. (2001, 1) write, “From this point of view, cognition depends on the kinds of experiences that come from having a body with particular perceptual and motor capabilities that are inseparably linked and that together form the matrix within which reasoning, memory, emotion, language, and all other aspects of mental life are meshed”. Seen in this way, fiction’s value is the way it helps us see thought as something that is produced through one’s interaction with an environment and not as an abstract process of reasoning in a vacuum. Understanding this relationship more precisely and how culturally specific it is is a promising area for future work.

## 5. Data Availability

Data can be found here: <https://doi.org/10.7910/DVN/KYL2N0>.

## 6. Software Availability

Code can be found here: <https://doi.org/10.7910/DVN/KYL2N0>.

## 7. Acknowledgements

This research was funded by the Social Sciences and Humanities Research Council of Canada (895-2013-1011).

## 8. Author Contributions

**Andrew Piper:** Conceptualization, Formal Analysis, Writing – original draft, Writing – review & editing.

## References

- Anderson, Amanda, Rita Felski, and Toril Moi (2019). *Character: Three Inquiries in Literary Studies*. University of Chicago Press.
- Bagga, Sunyam and Andrew Piper (2022). “HATHI 1M: Introducing a Million Page Historical Prose Dataset in English from the Hathi Trust”. In: *Journal of Open Humanities Data* 8. [10.5334/johd.71](https://doi.org/10.5334/johd.71).
- Bamman, David (2021). *BookNLP. A Natural Language Processing Pipeline for Books*. <https://github.com/booknlp/booknlp> (visited on 01/30/2022).
- Bamman, David, Sejal Popat, and Sheng Shen (2019). “An Annotated Dataset of Literary Entities”. In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, 2138–2144. [10.18653/v1/N19-1220](https://doi.org/10.18653/v1/N19-1220).
- Bamman, David, Ted Underwood, and Noah A Smith (2014). “A Bayesian Mixed Effects Model of Literary Character”. In: *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 370–379. [10.3115/v1/P14-1035](https://doi.org/10.3115/v1/P14-1035).
- Caracciolo, Marco and Karin Kukkonen (2021). *With Bodies: Narrative Theory and Embodied Cognition*. Ohio State University Press.
- Cheng, Jonathan (2020). “Fleshing out Models of Gender in English-language Novels (1850–2000)”. In: *Journal of Cultural Analytics* 5.1, 11652. [10.22148/001c.11652](https://doi.org/10.22148/001c.11652).
- Dunning, Ted E (1993). “Accurate Methods for the Statistics of Surprise and Coincidence”. In: *Computational linguistics* 19.1, 61–74. <https://aclanthology.org/J93-1003> (visited on 03/08/2024).
- Eder, Jens, Fotis Jannidis, and Ralf Schneider (2010). *Characters in Fictional Worlds*. De Gruyter.
- Frow, John (2014). *Character and Person*. Oxford University Press.

- Heuser, Ryan and Long Le-Khac (2012). *A Quantitative Literary History of 2,958 Nineteenth-Century British Novels: The Semantic Cohort Method*. Literary Lab Pamphlets. Stanford Literary Lab. <https://litlab.stanford.edu/LiteraryLabPamphlet4.pdf> (visited on 06/08/2023).
- Jannidis, Fotis (2004). *Figur und Person: Beitrag zu einer historischen Narratologie*. De Gruyter.
- Kidd, David, Martino Ongis, and Emanuele Castano (2016). "On Literary Fiction and its Effects on Theory of Mind". In: *Scientific Study of Literature* 6.1, 42–58. [10.1075/ssol.6.1.04kid](#).
- Palmer, Alan (2004). *Fictional Minds*. University of Nebraska Press.
- Phelan, James (1989). *Reading People, Reading Plots: Character, Progression, and the Interpretation of Narrative*. University of Chicago Press.
- Piper, Andrew (2018). *Enumerations: Data and Literary Study*. University of Chicago Press.
- (2022). "The CONLIT Dataset of Contemporary Literature". In: *Journal of Open Humanities Data* 8. [10.5334/johd.88](#).
- Thelen, Esther, Gregor Schöner, Christian Scheier, and Linda B Smith (2001). "The Dynamics of Embodiment: A Field Theory of Infant Perseverative Reaching". In: *Behavioral and Brain Sciences* 24.1, 1–34. [10.1017/s0140525x01003910](#).
- Underwood, Ted (2019). *Distant Horizons: Digital Evidence and Literary Change*. University of Chicago Press.
- Zunshine, Lisa (2006). *Why We Read Fiction: Theory of Mind and the Novel*. Ohio State University Press.