



Reading Literary Fiction Improves Theory of Mind

David Comer Kidd and Emanuele Castano

Science **342**, 377 (2013);

DOI: 10.1126/science.1239918

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines [here](#).

The following resources related to this article are available online at www.sciencemag.org (this information is current as of August 6, 2014):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/content/342/6156/377.full.html>

Supporting Online Material can be found at:

<http://www.sciencemag.org/content/suppl/2013/10/03/science.1239918.DC1.html>

This article **cites 27 articles**, 6 of which can be accessed free:

<http://www.sciencemag.org/content/342/6156/377.full.html#ref-list-1>

This article has been **cited by** 2 articles hosted by HighWire Press; see:

<http://www.sciencemag.org/content/342/6156/377.full.html#related-urls>

This article appears in the following **subject collections**:

Psychology

<http://www.sciencemag.org/cgi/collection/psychology>

tially toxic CNS waste products. Our analysis indicates that the cortical interstitial space increases by more than 60% during sleep, resulting in efficient convective clearance of A β and other compounds (Figs. 2 and 3). The purpose of sleep has been the subject of numerous theories since the time of the ancient Greek philosophers (34). An extension of the findings reported here is that the restorative function of sleep may be due to the switching of the brain into a functional state that facilitates the clearance of degradation products of neural activity that accumulate during wakefulness.

References and Notes

- C. B. Saper, P. M. Fuller, N. P. Pedersen, J. Lu, T. E. Scammell, *Neuron* **68**, 1023–1042 (2010).
- J. A. Hobson, *Nature* **437**, 1254–1256 (2005).
- B. A. Malow, *Epilepsy Curr.* **4**, 193–195 (2004).
- R. Stickgold, *Nature* **444**, 559–560 (2006).
- A. Rechtschaffen, M. A. Gilliland, B. M. Bergmann, J. B. Winter, *Science* **221**, 182–184 (1983).
- P. J. Shaw, G. Tononi, R. J. Greenspan, D. F. Robinson, *Nature* **417**, 287–291 (2002).
- P. Montagna, P. Gambetti, P. Cortelli, E. Lugaresi, *Lancet Neurol.* **2**, 167–176 (2003).
- J. R. Cirrito *et al.*, *Neuron* **48**, 913–922 (2005).
- M. E. Larson *et al.*, *J. Neurosci.* **32**, 10253–10266 (2012).
- K. Yamada *et al.*, *J. Neurosci.* **31**, 13110–13117 (2011).
- K. Aukland, R. K. Reed, *Physiol. Rev.* **73**, 1–78 (1993).
- J. J. Iloff *et al.*, *Sci. Transl. Med.* **4**, 147ra111 (2012).
- J. J. Iloff *et al.*, *J. Clin. Invest.* **123**, 1299–1309 (2013).
- M. Nedergaard, *Science* **340**, 1529–1530 (2013).
- R. J. Bateman *et al.*, *Nat. Med.* **12**, 856–861 (2006).
- J. E. Kang *et al.*, *Science* **326**, 1005–1007 (2009).
- C. Nicholson, J. M. Phillips, *J. Physiol.* **321**, 225–257 (1981).
- C. Nicholson, *J. Neurosci. Methods* **48**, 199–213 (1993).
- X. Yao, S. Hrabětová, C. Nicholson, G. T. Manley, *J. Neurosci.* **28**, 5460–5464 (2008).
- C. Nicholson, E. Syková, *Trends Neurosci.* **21**, 207–215 (1998).
- E. Syková, C. Nicholson, *Physiol. Rev.* **88**, 1277–1340 (2008).
- E. Syková *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 479–484 (2005).
- J. P. Kinney *et al.*, *J. Comp. Neurol.* **521**, 448–464 (2013).
- R. Deane *et al.*, *J. Clin. Invest.* **118**, 4002–4013 (2008).
- M. Steriade, D. A. McCormick, T. J. Sejnowski, *Science* **262**, 679–685 (1993).
- M. E. Carter *et al.*, *Nat. Neurosci.* **13**, 1526–1533 (2010).
- C. M. Constantinople, R. M. Bruno, *Neuron* **69**, 1061–1068 (2011).
- J. O'Donnell, D. Zeppenfeld, E. McConnell, S. Pena, M. Nedergaard, *Neurochem. Res.* **37**, 2496–2512 (2012).
- C. J. McBain, S. F. Traynelis, R. Dingledine, *Science* **249**, 674–677 (1990).
- K. Parameshwaran, M. Dhanasekaran, V. Suppiramaniam, *Exp. Neurol.* **210**, 7–13 (2008).
- K. V. Kuchibhotla *et al.*, *Neuron* **59**, 214–225 (2008).
- M. P. Mattson, *Ann. N. Y. Acad. Sci.* **747**, 50–76 (1994).
- A. Borbely, I. Tobler, in *Brain Mechanisms of Sleep*, D. J. McGinty, Ed. (Raven, New York, 1985), pp. 35–44.
- J. Barbera, *Sleep Med.* **9**, 906–910 (2008).

Acknowledgments: This study was supported by NIH/National Institute of Neurological Disorders and Stroke (NS078167 and NS078304 to M.N. and NS028642 to C.N.). We thank S. Veasey for comments on the manuscript.

Supplementary Materials

www.sciencemag.org/content/342/6156/373/suppl/DC1
Materials and Methods
Figs. S1 to S5
References

30 May 2013; accepted 28 August 2013
10.1126/science.1241224

Reading Literary Fiction Improves Theory of Mind

David Comer Kidd* and Emanuele Castano*

Understanding others' mental states is a crucial skill that enables the complex social relationships that characterize human societies. Yet little research has investigated what fosters this skill, which is known as Theory of Mind (ToM), in adults. We present five experiments showing that reading literary fiction led to better performance on tests of affective ToM (experiments 1 to 5) and cognitive ToM (experiments 4 and 5) compared with reading nonfiction (experiments 1), popular fiction (experiments 2 to 5), or nothing at all (experiments 2 and 5). Specifically, these results show that reading literary fiction temporarily enhances ToM. More broadly, they suggest that ToM may be influenced by engagement with works of art.

The capacity to identify and understand others' subjective states is one of the most stunning products of human evolution. It allows successful navigation of complex social relationships and helps to support the empathic responses that maintain them (1–5). Deficits in this set of abilities, commonly referred to as Theory of Mind (ToM), are associated with psychopathologies marked by interpersonal difficulties (6–8). Even when the ability is intact, disengagement of ToM has been linked to the breakdown of positive interpersonal and intergroup relationships (9).

Researchers have distinguished between affective ToM (the ability to detect and understand others' emotions) and cognitive ToM (the inference and representation of others' beliefs and in-

tentions) (7, 8). The affective component of ToM, in particular, is linked to empathy (positively) and antisocial behavior (negatively) (7, 8). It is thus not surprising that we foster ToM in our children by having them attend to the emotional states of others: “Do you think he is happy or sad as a consequence of your action?” Such explicit encouragements to understand others usually diminish when children appear to skillfully and empathically engage in interpersonal relationships. Cultural practices, though, may function to promote and refine interpersonal sensitivity throughout our lives. One such practice is reading fiction.

Familiarity with fiction, self-reported empathy, and performance on an advanced affective ToM test have been correlated (10, 11), and limited experimental evidence suggests that reading fiction increases self-reported empathy (12, 13). Fiction seems also to expand our knowledge of others' lives, helping us recognize our similarity to them (10, 11, 14). Although fiction

may explicitly convey social values and reduce the strangeness of others, the observed relation between familiarity with fiction and ToM may be due to more subtle characteristics of the text. That is, fiction may change how, not just what, people think about others (10, 11, 14). We submit that fiction affects ToM processes because it forces us to engage in mind-reading and character construction. Not any kind of fiction achieves that, though. Our proposal is that it is literary fiction that forces the reader to engage in ToM processes.

The category of literary fiction has been contested on the grounds that it is merely a marker of social class, but features of the modern literary novel set it apart from most best-selling thrillers or romances. Miall and Kuiken (15–17) emphasize that through the systematic use of phonological, grammatical, and semantic stylistic devices, literary fiction defamiliarizes its readers. The capacity of literary fiction to unsettle readers' expectations and challenge their thinking is also reflected in Roland Barthes's (18) distinction between writerly and readerly texts. Although readerly texts—such as most popular genre fiction—are intended to entertain their mostly passive readers, writerly—or literary—texts engage their readers creatively as writers. Similarly, Mikhail Bakhtin (19) defined literary fiction as polyphonic and proposed that readers of literary fiction must contribute their own to a cacophony of voices. The absence of a single authorial perspective prompts readers to enter a vibrant discourse with the author and her characters.

Bruner (20), like Barthes and Bakhtin, has proposed that literature engages readers in a discourse that forces them to fill in gaps and search “for meanings among a spectrum of possible meanings” (p. 25). Bruner argues that to elicit

The New School for Social Research, 80 Fifth Avenue, New York, NY 10011, USA.

*Corresponding author. E-mail: kidd305@newschool.edu (D.C.K.); castano@newschool.edu (E.C.)

this writerly stance, literary fiction triggers presupposition (a focus on implicit meanings), subjectification [depicting reality “through the filter of the consciousness of protagonists in the story” (p. 25)], and multiple perspectives (perceiving the world simultaneously from different viewpoints). These features mimic those of ToM.

Our contention is that literary fiction, which we consider to be both writerly and polyphonic, uniquely engages the psychological processes needed to gain access to characters’ subjective experiences. Just as in real life, the worlds of literary fiction are replete with complicated individuals whose inner lives are rarely easily discerned but warrant exploration. The worlds of fiction, though, pose fewer risks than the real world, and they present opportunities to consider the experiences of others without facing the potentially threatening consequences of that engagement. More critically, whereas many of our mundane social experiences may be scripted by convention and informed by stereotypes, those presented in literary fiction often disrupt our expectations. Readers of literary fiction must draw on more flexible interpretive resources to infer the feelings and thoughts of characters. That is, they must engage ToM processes. Contrary to literary fiction, popular fiction, which is more readerly, tends to portray the world and characters as internally consistent and predictable (21). Therefore, it may reaffirm readers’ expectations and so not promote ToM.

To test our general hypothesis that literary fiction would prime ToM, we first compared the effects of reading literary fiction with reading nonfiction (experiment 1) and then focused on testing our predictions about the different effects of reading literary and popular fiction (experiments 2 to 5).

Difficulty in precisely quantifying literariness notwithstanding, some works are considered particularly good examples of literature and are recognized with prestigious awards (e.g., the National Book Award). Although selected through an inherently inexact process, prize-winning texts are more likely to embody general characteristics of literature than bestsellers of genre fiction (e.g., romance and adventure stories). In the absence of a clear means of quantifying literariness, the judgments of expert raters (i.e., literary prize jurors) were used. Accordingly, to study the effects of reading literary fiction, we selected literary works of fiction by award-winning or canonical writers and compared their effects on ToM with reading nonfiction, popular fiction, or nothing at all.

In experiment 1 (22), 86 participants were randomly assigned to read one of six short texts (three literary fiction and three nonfiction). Next, participants completed a false-belief test as a measure of cognitive ToM (23) and an advanced affective ToM test, the Reading the Mind in the Eyes Test [RMET (6)], in which they were asked to identify facially expressed emotions. Participants’ familiarity with fiction was assessed using the Author Recognition Test (24), an index of general exposure

to fiction that avoids problems of socially desirable responding. Affect (25), engagement with the text (transportation scale) (26), and demographic information were assessed.

For the cognitive ToM task, participants were asked to indicate the probability that a character would act according to the character’s own false belief or the participant’s true belief. Participants ($n = 13$) who failed to give probabilities and univariate outliers (>3.5 SD from the mean; $n = 6$) were excluded from the analysis. Probabilities were compared in a 2(false-belief versus no false-belief condition) \times 2(fiction versus nonfiction) analysis of variance (ANOVA). There was no main effect for the type of scenario, which suggests no evidence of egocentric bias ($F_{1,63} = 1.47$, $P = 0.22$). The level of false estimates was low across conditions (grand mean \pm standard deviation, 6.61 ± 9.79).

Scores for the affective ToM task were computed by summing the number of correct identifications of facially expressed emotions (6) and analyzed using ANOVA, with condition and Author Recognition Test as between-participants factors (Table 1). Scores were higher in the literary fiction than nonfiction condition (Table 2). Higher Author Recognition Test scores (indicating more familiarity with fiction) predicted higher RMET scores. When entered as covariates, education; gender; age; transportation; negative affect; self-reported sadness; and average time spent on RMET items did not significantly alter the main effect of condition ($P = 0.05$). More time spent on RMET items predicted better performance ($\beta = 0.23$, $P = 0.02$). No other covariates approached significance (P values of >0.14).

Experiment 2 aimed to replicate and extend the findings of experiment 1 by using different texts and a different measure of affective ToM, the Diagnostic Analysis of Nonverbal Accuracy 2—Adult Faces test (DANVA2-AF) (27). Experiment 2 was also designed to directly differentiate between the effects of popular versus literary fiction (28).

Participants ($n = 114$) were randomly assigned to read one of three excerpts from recent

finalists for the National Book Award (literary fiction condition), one of three excerpts from recent bestsellers on Amazon.com (popular fiction condition), or nothing at all (no-reading condition) (22). Participants then completed the measure of cognitive ToM used in experiment 1 and the DANVA2-AF before completing the Author Recognition Test, the transportation scale, and demographic questions. Performance on the false-belief cognitive ToM task was analyzed as in experiment 1, but no significant effects were detected (P values of >0.13).

DANVA2-AF scores were computed by summing errors on all of the negative affect items (22). Untransformed means are reported, but log-transformed scores were used in an ANOVA with experimental condition and Author Recognition Test as between-participants factors (see Table 1). No interaction emerged, but higher scores on the Author Recognition Test were weakly associated with fewer errors on the DANVA2-AF. The omnibus main effect of condition was marginally significant, and the pairwise comparisons revealed significant differences between conditions consistent with our hypothesis. Fewer errors were made in the literary fiction condition than in the no-reading and popular fiction conditions, whereas there was no difference between the latter two ($P = 0.98$) (see Table 2). As in experiment 1, education; gender; and age were not significant covariates (P values of >0.34) and did not alter the critical, omnibus main effect of condition ($P = 0.08$). Transportation did not correlate with DANVA2-AF scores ($P = 0.94$).

Experiment 3 ($N = 69$) aimed to replicate the literary fiction versus popular fiction comparison (22). The popular fiction texts were three stories from an edited anthology of popular fiction (29), and literary fiction texts were three stories from a collection of the 2012 PEN/O. Henry Award winners for short stories (30). Participants’ affect was assessed using the Positive Affect Negative Affect Scale (PANAS) and a single-item report of sadness. Using the same analytical strategy used in experiment 1, it was found that RMET scores

Table 1. RMET and DANVA2-AF analyses.

Experiment	Independent variable	Test	<i>P</i>	ω_p^2
Exp. 1 RMET	Condition	$F_{1,82} = 6.40$	0.01	0.05
	Author Recognition Test	$\beta = 0.36$	0.0003	0.13
	Author Recognition Test \times Condition	$F_{1,82} = 1.06$	0.30	0.00
Exp. 2 DANVA2-AF	Condition	$F_{2,108} = 2.57$	0.08	0.02
	Author Recognition Test	$\beta = -0.16$	0.08	0.01
	Author Recognition Test \times Condition	$F_{2,108} = 1.17$	0.31	0.00
Exp. 3 RMET	Condition	$F_{1,65} = 4.07$	0.04	0.04
	Author Recognition Test	$\beta = -0.01$	0.90	-0.01
Exp. 4 RMET	Author Recognition Test \times Condition	$F_{1,65} = 0.01$	0.90	-0.01
	Condition	$F_{1,68} = 4.39$	0.04	0.04
	Author Recognition Test	$\beta = 0.39$	<0.001	0.15
Exp. 5 RMET	Author Recognition Test \times Condition	$F_{1,68} = 1.50$	0.22	0.00
	Condition	$F_{2,352} = 3.10$	0.04	0.01
	Author Recognition Test	$\beta = 0.28$	<0.001	0.07
	Author Recognition Test \times Condition	$F_{2,352} = 1.37$	0.25	0.00

were higher in the literary fiction condition than in the popular fiction condition. There were no effects involving the Author Recognition Test (for test, Table 1; for means, Table 2). Education, gender, and the average time spent on RMET items were not significant covariates (P values of >0.12) and did not alter the effect of condition ($P = 0.04$).

In experiments 1 and 2, no effects were observed on the cognitive ToM measure, a false-belief task. Since participants in neither condition clearly failed to recruit cognitive ToM, it is possible that the task may have been insufficiently sensitive. Therefore a fourth experiment included the Yoni test (7). The Yoni test is a new measure that has been used in only a handful of studies. However, it has been validated (7, 8, 31) and has the advantage of assessing both cognitive and affective ToM.

In experiment 4, four of the texts used in experiment 3 along with two new stories, one for each condition (i.e., literary fiction and popular fiction), from the same sources were used (22). Participants ($N = 72$) completed the RMET and the Yoni test. For the 24 cognitive and 24 affective ToM trials in the Yoni test, participants must draw from minimal linguistic and visual cues to infer a character's thoughts and emotions, respectively. An additional 16 control trials require the identification of spatial relations. For each type of item, there are equal numbers of trials requiring first-order and second-order (more difficult) inferences.

RMET scores were higher in the literary fiction condition than in the popular fiction condition (for tests, see Table 1; for means, see Table 2). Author Recognition Test scores predicted RMET scores. Entered as covariates, subject variables (i.e., education, age, and gender) did not reach significance (P values of >0.14), though time spent on RMET items did ($\beta = 0.21$, $P = 0.04$). However, the effect of condition was only slightly altered and remained significant ($P = 0.05$).

Yoni performance was analyzed via a mixed analysis of covariance (ANCOVA) with type (affective versus cognitive) and level of difficulty (first order versus second order) of trials as within-participants factors, condition and Author Recognition Test scores as between-participants factors,

and scores on the control task as a covariate (31). A main effect of condition emerged ($F_{1,67} = 4.47$, $P = 0.03$, $\omega_p^2 = 0.04$) but no other effects involving condition or Author Recognition Test scores approached significance (P values of >0.27). Other significant effects, which are not relevant to the hypotheses, are described in the supplementary materials (22). Participants in the literary fiction condition [0.89 ± 0.08 , 95% confidence interval (CI) = 0.86, 0.92] performed with greater accuracy on all ToM trials than those in the popular fiction condition (0.85 ± 0.10 , CI = 0.82, 0.87).

A fifth experiment (22) aimed to replicate experiment 4 and test for the influences of subject variables (i.e., education, age, gender) and possible confounds with a larger sample ($N = 356$). As in experiments 3 and 4, three works of literary fiction were taken from a collection of the 2012 PEN/O. Henry Prize winners (30) and three works of popular fiction from an anthology (29). Participants were randomly assigned to the literary fiction, popular fiction, or no-reading control condition; completed the RMET and Yoni tasks; reported their current affect (PANAS), along with two additional items assessing sadness and happiness; and completed the Author Recognition Test. Participants in the two reading conditions completed the transportation scale and two additional items assessing the extent to which they enjoyed reading the text and how much they thought it represented "excellent literature." All participants reported their age, gender, ethnicity, and highest level of attained education before being debriefed and compensated.

Literary texts (3.54 ± 1.31 , CI = 3.28, 3.80) were enjoyed less than popular texts (4.07 ± 1.53 , CI = 3.80, 4.34; $F_{1,223} = 7.62$, $P = 0.006$, $\omega_p^2 = 0.02$), but they were seen as better examples of literature (4.84 ± 1.40 , CI = 4.56, 5.11) than popular texts (4.43 ± 1.60 , CI = 4.15, 4.72; $F_{1,223} = 4.04$, $P = 0.04$, $\omega_p^2 = 0.01$). Reported transportation did not significantly differ across conditions ($F_{1,223} = 3.20$, $P = 0.07$, $\omega_p^2 = 0.00$), although it was slightly higher in the literary condition (3.90 ± 0.41 , CI = 3.83, 3.98) than the popular condition (3.81 ± 0.39 , CI = 3.73, 3.88). None of these variables were correlated (P values of

>0.11) with performance on either the RMET or the Yoni task (controlling for performance on physical trials).

Results on the RMET were analyzed as in the previous experiments. The effect of condition was significant (see Table 1). Scores were significantly higher in the literary fiction condition than in the popular fiction and no-reading conditions (see Table 2). The latter two conditions did not differ ($P = 0.65$). A significant main effect of Author Recognition Test scores emerged (see Table 1). Added as covariates, gender, education, age, positive affect, negative affect, sadness, happiness, and time spent on RMET items did not significantly relate to RMET scores (P values of >0.23), and the effect of condition was only slightly altered ($P = 0.06$).

The analytical strategy used in experiment 4 was also used for the Yoni task. The main effect of condition ($F_{2,351} = 0.64$, $P = 0.52$) was not significant, but there was a significant interaction of condition and the two within-subjects factors, trial difficulty and trial type ($F_{2,351} = 3.42$, $P = 0.03$). The interaction of Author Recognition Test scores, trial difficulty, and trial type approached significance ($F_{1,351} = 2.88$, $P = 0.09$), but no other effects involving condition or Author Recognition Test did (P values of >0.11). Other significant effects, which are not relevant to the hypotheses, are described in the supplementary materials (22). To disentangle the three-way interaction including the experimental condition, a repeated measures ANCOVA, with item type (cognitive, affective) and condition as factors, and performance on the control task as covariate, was conducted separately for first-order and second-order trials. On first-order trials, there was a main effect of the covariate ($\beta = 0.20$, $P < 0.001$, $\omega_p^2 = 0.03$) and of condition ($F_{2,351} = 4.21$, $P = 0.01$, $\omega_p^2 = 0.01$). No other effects approached significance (P values of >0.87). Pair-wise comparisons revealed that scores were higher in the literary fiction condition (0.98 ± 0.02 , CI = 0.97, 0.99) than in the popular fiction condition (0.96 ± 0.06 , CI = 0.95, 0.97; $t = 2.85$, $P = 0.004$) and the no-reading condition (0.97 ± 0.05 , CI = 0.96, 0.98; $t = 2.01$, $P = 0.04$). The popular fiction condition and no-reading condition did not differ ($P = 0.33$). On second-order trials, no effects involving condition or Author Recognition Test scores approached significance (P values of >0.16).

The difference between first- and second-order trials, which appeared only in experiment 5, might be due to its higher statistical power, which allowed for this difference to be detected. The second-order Yoni trials may require a set of more advanced cognitive skills (e.g., metarepresentation) that are less easily influenced by manipulation than the other tasks, all of which are first-order ToM tasks.

Experiment 1 showed that reading literary fiction, relative to nonfiction, improves performance on an affective ToM task. Experiments 2 to 5 showed that this effect is specific to literary fiction. On

Table 2. Means (adjusted for other terms in the models) and standard deviations of RMET and DANVA2-AF scores. 95% confidence intervals are reported in brackets. X, no data. Means in the same row that share the same superscripts differ at $P < 0.05$.

Experiment	Literary fiction	Popular fiction	No reading	Nonfiction
Exp. 1 RMET	25.90 ± 4.38 ^a [24.55, 27.24]	X	X	23.47 ± 5.17 ^a [22.13, 24.82]
Exp. 2 DANVA2-AF	4.70 ± 2.31 ^{ab} [3.79, 5.61]	5.85 ± 2.93 ^a [4.96, 6.74]	5.86 ± 2.89 ^b [5.00, 6.72]	X
Exp. 3 RMET	25.92 ± 4.07 ^a [23.99, 27.86]	23.22 ± 6.16 ^a [21.34, 25.09]	X	X
Exp. 4 RMET	26.19 ± 5.43 ^a [24.52, 27.85]	23.71 ± 5.08 ^a [22.18, 25.24]	X	X
Exp. 5 RMET	26.21 ± 3.59 ^{ab} [25.45, 26.97]	24.96 ± 4.60 ^a [24.18, 25.74]	25.20 ± 4.69 ^b [24.99, 25.91]	X

cognitive measures, no effects emerged on the false-belief task used in experiments 1 and 2. Because error rates on the false-belief task were very low, the measure may have been insufficiently sensitive to capture the effects of the manipulations. However, on the more-demanding Yoni task used in experiments 4 and 5, the effect on cognitive trials was present and indistinguishable from that on affective trials.

The Author Recognition Test predicted RMET scores in experiments 1, 4, and 5, and success on the DANVA2-AF (marginally) in experiment 2; but it did not predict performance on the Yoni task or, anomalously, the RMET in experiment 3. Thus, although generally consistent with previous findings (10–12), our pattern of results suggests the need for further research into the relation between measures of familiarity with fiction and performance on different ToM tasks.

The results of five experiments support our hypothesis that reading literary fiction enhances ToM. Existing explanations focused on the content of fiction cannot account for these results. First, the texts we used varied widely in subject matter. Second, it is unlikely that people learned much more about others by reading any of the short texts. Third, the effects were specific to literary fiction. We propose that by prompting readers to take an active writerly role to form representations of characters' subjective states, literary fiction recruits ToM. The evidence we report here is consistent with this view, but we see these findings as preliminary and much research is needed.

First, our findings demonstrate the short-term effects of reading literary fiction. However, taken together, the relation between the Author Recognition Test and ToM performance and the finding that it is specifically literary fiction that facilitates ToM processes suggest that reading literary fiction may lead to stable improvements in ToM. Because the Author Recognition Test does not distinguish between exposure to literary and popular fiction, additional research with refined methods is necessary to test this hypothesis.

Second, literary fiction, like many stimuli drawn from the real world, is heterogeneous and complex. Although it is not clearly quantifiable, literariness has ecological validity as a construct, as suggested by participants' agreement with prize jurors on the literariness of the texts in experiment 5. On the basis of strategies used by researchers studying violent video games [e.g., (32)] and fiction (12), literariness was held relatively constant in each condition while potentially confounding features varied. Self-reported affect along with transportation into, enjoyment, and perceived literariness of the texts did not account for the effects of condition. Further analyses tested the roles of superficial linguistic features of the texts. Frequencies of negative and positive emotion terms, social words, cognitive words, big words (more than six letters), and self-references were computed in each text using Linguistic Inquiry and Word Count (LIWC) software (33). Stan-

dardized RMET or DANVA2-AF scores from all experiments were analyzed using ANCOVA, with experimental condition and Author Recognition Test scores as factors and all six LIWC variables as covariates (data from the no-reading conditions was not included). The frequency of negative emotion words ($\beta = 0.09$, $P = 0.05$, $\omega_p^2 = 0.00$) positively predicted ToM scores, but no other effects of LIWC variables approached significance (P values of >0.17). The main effects of condition ($F_{1,515} = 12.02$, $P < 0.001$, $\omega_p^2 = 0.02$) and Author Recognition Test scores ($\beta = 0.23$, $P < 0.001$, $\omega_p^2 = 0.05$) remained significant. This result suggests that the effect of literature observed across experiments may not be easily reduced to superficial linguistic characteristics. Future research, notably following the lead of Miall and Kuiken (15–17), as well as Bruner (20), may reveal more subtle, but nonetheless quantifiable, features that set literary fiction apart.

The present findings mark only one step toward understanding the impact of our interactions with fiction, the experiences of which are thought to contribute to the development of consciousness and to enrich our daily lives (34). Indeed, there are surely many consequences of reading on cognitive and affective processes that are independent of its effects on ToM, and it seems likely that many of those may result from popular, as well as literary, fiction. Similarly, whereas literary fiction appears able to promote ToM, this capacity does not fully capture the concept of literariness, which includes, among others, aesthetic and stylistic matters not addressed in this research. It is our hope that further research will focus on other forms of art, such as plays and movies, that involve identifying and interpreting the subjective experiences of others (10, 28).

Literature has been deployed in programs intended to promote social welfare, such as those intended to promote empathy among doctors (35) and life skills among prisoners (36). Literature is, of course, also a required subject throughout secondary education in the United States, but reformers have questioned its importance: A new set of education standards that has been adopted by 46 U.S. states (the Common Core State Standards) controversially calls for less emphasis on fiction in secondary education [see (37)]. Debates over the social value of types of fiction and the arts more broadly are important, and it seems critical to supplement them with empirical research. These results show that reading literary fiction may hone adults' ToM, a complex and critical social capacity.

References and Notes

1. C. Batson, in *The Handbook of Social Psychology*, D.T. Gilbert, S.T. Fiske, G. Lindzey, Eds. (McGraw-Hill, Boston, 1998), pp. 282–316.
2. F. B. de Waal, *Science* **336**, 874–876 (2012).
3. R. Saxe, S. Carey, N. Kanwisher, *Annu. Rev. Psychol.* **55**, 87–124 (2004).
4. C. L. Sebastian et al., *Soc. Cogn. Affect. Neurosci.* **7**, 53–63 (2012).
5. S. G. Shamay-Tsoory, *Neuroscientist* **17**, 18–24 (2011).

6. S. Baron-Cohen, S. Wheelwright, J. Hill, Y. Raste, I. Plumb, *J. Child Psychol. Psychiatry* **42**, 241–251 (2001).
7. S. G. Shamay-Tsoory, J. Aharon-Peretz, *Neuropsychologia* **45**, 3054–3067 (2007).
8. S. G. Shamay-Tsoory, H. Harari, J. Aharon-Peretz, Y. Levkovitz, *Cortex* **46**, 668–677 (2010).
9. E. Castano, in *The Oxford Handbook of Personality and Social Psychology*, K. Deaux and M. Snyder, Eds. (Oxford Univ. Press, New York, 2012), pp. 419–445.
10. R. Mar, K. Oatley, J. Hirsh, J. dela Paz, J. Peterson, *J. Res. Pers.* **40**, 694–712 (2006).
11. R. Mar, K. Oatley, J. Peterson, *Communications* **34**, 407–428 (2009).
12. M. Djikic, K. Oatley, M. Moldoveanu, *Scientific Study of Literature* **3**, 28–47 (2013).
13. P. M. Bal, M. Veltkamp, *PLoS ONE* **8**, e55341 (2013).
14. E. Schiappa, P. Gregg, D. Hewes, *Commun. Monogr.* **72**, 92–115 (2005).
15. D. Miall, D. Kuiken, *Poetics* **22**, 389–407 (1994).
16. D. Miall, D. Kuiken, *Discourse Process.* **17**, 337–352 (1994).
17. D. Miall, D. Kuiken, *Discourse Process.* **28**, 121–138 (1999).
18. R. Barthes, *SZ: An Essay* (Hill and Wang, New York, 1974).
19. M. Bakhtin, *Problems of Dostoevsky's Poetics* (Univ. Minnesota, Minneapolis, 1984).
20. J. Bruner, *Actual Minds, Possible Worlds* (Harvard Univ., Cambridge, MA, 1986).
21. R. Gerrig, D. Rapp, *Poetics Today* **25**, 265–281 (2004).
22. Materials and methods are available as supplementary materials on Science Online.
23. B. A. Converse, S. Lin, B. Keyser, N. Epley, *Emotion* **8**, 725–730 (2008).
24. D. J. Acheson, J. B. Wells, M. C. MacDonald, *Behav. Res. Methods* **40**, 278–289 (2008).
25. D. Watson, L. A. Clark, A. Tellegen, *J. Pers. Soc. Psychol.* **54**, 1063–1070 (1988).
26. M. C. Green, T. C. Brock, *J. Pers. Soc. Psychol.* **79**, 701–721 (2000).
27. S. Nowicki, *Manual for the Receptive Tests of the Diagnostic Analysis of Nonverbal Accuracy 2* (Department of Psychology, Emory Univ., Atlanta, GA, 2010).
28. R. Mar, K. Oatley, *Perspect. Psychol. Sci.* **3**, 173–192 (2008).
29. G. Hoppenstand, Ed., *Popular Fiction: An Anthology* (Longman, New York, 1998).
30. L. Furman, Ed., *The PEN/O. Henry Prize Stories 2012: The Best Stories of the Year* (Anchor, New York, 2012).
31. S. G. Shamay-Tsoory, *J. Autism Dev. Disord.* **38**, 1451–1461 (2008).
32. B. J. Bushman, C. A. Anderson, *Psychol. Sci.* **20**, 273–277 (2009).
33. J. Pennebaker, R. Booth, M. Francis, *Linguistic Inquiry and Word Count (LIWC): LIWC2007* (LIWC.net, Austin, TX, 2007).
34. N. Humphrey, *Consciousness Regained* (Oxford Univ., Oxford, 1983).
35. M. F. McLellan, A. H. Jones, *Lancet* **348**, 109–111 (1996).
36. J. Billington, *Critical Survey* **23**, 67–85 (2011).
37. S. Mosle, *New York Times*, 22 November 2012; <http://opinionator.blogs.nytimes.com/2012/11/22/what-should-children-read/?r=1>.

Acknowledgments: Data are available in spreadsheet format from authors by request. This research was funded by a Prize Fellowship and a Dissertation Fellowship by the New School for Social Research to the first author, and a Faculty Development Award by the New School for Social Research to the second author. We are grateful to Andrew Marzoni for his insight into literary criticism and pointing us toward Barthes's *SZ* (19).

Supplementary Materials

www.sciencemag.org/content/ful/342/6156/PAGE/suppl/DC1
Materials and Methods
Supplementary Text
References (38–51)

1 May 2013; accepted 18 September 2013
Published online 3 October 2013;
10.1126/science.1239918