Journal of Experimental Psychology: General
2018, Vol. 147, No. 11, 1713–1727

0096-3445/18/$12.00 http://dx.doi.org/10.1037/xge0000395

Fiction Reading Has a Small Positive Impact on Social Cognition: A Meta-Analysis

David Dodell-Feder
McLean Hospital, Belmont, Massachusetts, and
Harvard Medical School

Diana I. Tamir
Princeton University

Scholars from both the social sciences and the humanities have credited fiction reading with a range of positive real-world social effects. Research in psychology has suggested that readers may make good citizens because fiction reading is associated with better social cognition. But does fiction reading *causally* improve social cognition? Here, we meta-analyze extant published and unpublished experimental data to address this question. Multilevel random-effects meta-analysis of 53 effect sizes from 14 studies demonstrated that it does: compared to nonfiction reading and no reading, fiction reading leads to a small, statistically significant improvement in social-cognitive performance (g = .15–.16). This effect is robust across sensitivity analyses and does not appear to be the result of publication bias. We recommend that in future work, researchers use more robust reading manipulations, assess whether the effects transfer to improved real-world social functioning, and investigate mechanisms.

Keywords: fiction, reading, social cognition, theory of mind, empathy

Supplemental materials: http://dx.doi.org/10.1037/xge0000395.supp

Open a work of literary fiction and you immediately gain access to the inner workings of another person's mind. This is a remarkable phenomenon: from a scant set of words, readers can construct nuanced mental lives of story characters. Reading fiction does more than just open a window into the minds of *fictional* characters; reading fiction may also help readers navigate the *real* social world. Indeed, scholars have credited fiction reading with societal shifts toward civility (Pinker, 2011). Research in psychology has suggested that readers make good citizens because reading may improve one's social-cognitive ability (Mar & Oatley, 2008; Oatley, 2016); that is, one's ability to perceive, interpret, and respond to social information (Fiske & Taylor, 2013). A meta-analysis of correlational studies, for instance, has shown that frequent readers of fiction score higher on measures of empathy and theory of mind—the ability to think about others' minds—than nonreaders (Mumper & Gerrig, 2017).

However, these correlational findings leave open the possibility that people with better social skills simply read more fiction, or that a third confounding variable influences both. Does fiction reading causally improve social cognition? There are theoretical and empirical reasons to believe it does. Researchers have argued that fiction may impact social cognition for two reasons (Mar, 2015; Oatley, 2016). First, fiction may induce the *process* of simulating story characters—including their social, mental, and emotional experiences. In this way, readers may get extra practice with the same social processes they engage during real-world social cognition (Oatley, 2016). The notion that fiction reading and social cognition engage similar processes is further supported by neuroimaging work demonstrating an overlap in the networks recruited during story reading and theory of mind (Mar, 2011), and increased engagement of the brain's default mode network while simulating literary passages with social content (Tamir, Bricker, Dodell-Feder, & Mitchell, 2016). Second, fiction may provide concrete *content* about human psychology and social interaction, and about distant countries, cultures, and peoples that readers may never have access to otherwise (Mar & Oatley, 2008). In this sense, fiction may help to build up a reader's social knowledge.

Empirically, some recent experimental studies have suggested that fiction reading indeed causally improves people's social cognition. In these studies, people randomly assigned to read short fictional stories, excerpts, or books, outperform people who are asked to read nonfiction or nothing on a variety of social-cognitive tasks (Black & Barnes, 2015; Kidd & Castano, 2013; Kidd, Ongis,

This article was published Online First February 26, 2018.

David Dodell-Feder, Institute for Technology in Psychiatry, McLean Hospital, Belmont, Massachusetts, and Department of Psychiatry, Harvard Medical School; Diana I. Tamir, Department of Psychology, Princeton Neuroscience Institute, Princeton University.

David Dodell-Feder's is now with the Department of Psychology at the University of Rochester.

We thank the researchers who shared data and information—Matthijs Bal, Jessica Black, Maja Djikic, Dan Johnson, David Kidd, Emy Koopman, Raymond Mar, Micah L. Mumper, M. E. Panero, Dalya Samur, Deena Weisberg—and Steven Worthington at the Institute for Quantitative Social Sciences at Harvard University for assistance with data analysis. Both authors developed the study concept. D. Dodell-Feder collected and analyzed the data under the supervision of D. Tamir. D. Dodell-Feder wrote the first draft of the manuscript, and D. Tamir critically revised the manuscript. Both authors approved the final version of the manuscript for submission.

Correspondence concerning this article should be addressed to David Dodell-Feder, Department of Psychology, University of Rochester, 453 Meliora Hall, Rochester, NY 14627. E-mail: d.dodell-feder@rochester.edu

& Castano, 2016; Pino & Mazza, 2016). However, several recent nonreplications of these findings (Panero et al., 2016; Samur, Tops, & Koole, 2017) cast doubt on whether the effect is robust. If reading just a short work of fiction improves social cognition, the implications would be far-reaching on both an individual and societal level. However, the field should know if fiction reading does *not* actually impact social cognition before investing additional empirical and societal resources in understanding and applying this effect.

In order to help resolve this debate, here we analyze the extant published and unpublished data from experimental investigations into fiction's impact on social cognition using robust meta-analytic methods. Our primary aim is to evaluate whether, and to what extent, fiction reading improves social cognition. Social cognition includes a broad suite of abilities related to processing, interpreting, and responding to social information. Prior work in this field has refrained from making specific hypotheses regarding the aspects of social cognition fiction may improve. As such, here, we analyze fiction's effect on a range of social-cognitive tasks, including theory of mind, empathy (or emotion sharing), and prosocial behaviors. As a secondary aim, we assess which of several factors may influence whether, or to what extent fiction reading influences social–cognitive ability, including how social cognition is measured, and what reading fiction is compared to (e.g., nonfiction, no reading).

Method

Study Search and Retrieval

We conducted the literature search through four parallel routes. First, we queried PubMed, PsychInfo, and Web of Science for material from academic journals, dissertations, conference proceedings, and editorials. We limited results to English, but did not limit the date of publication. In line with other work on social–cognitive assessments (Pinkham, Penn, Green, & Harvey, 2015), and for the sake of thoroughness, we used 10 search terms related to social cognition in addition to a term for fiction: *fiction* AND (*social cognition* OR *social ability* OR *social skill* OR *social perception* OR *theory of mind* OR *mentalizing* OR *mind reading* OR *perspective taking* OR *empath** OR *emotion*). We also included common variants of these terms (e.g., *theory of mind* and *theory-of-mind*). We initially conducted the search in February 2016, and then performed the same search again in August 2016 to keep our records up-to-date. Second, we posted calls for unpublished data on the Society for Personality and Social Psychology (SPSP) listserv and the International Society for the Empirical Study of Literature and Media (IGEL) listserv. Third, after performing an initial screening of the online database search, we contacted the authors of the papers to be included in the analysis requesting unpublished data, including conference proceedings. Finally, we reviewed the reference list of relevant review papers as well as papers to be included in the analysis for other sources. The search returned 751 nonduplicated records, eight of which were from nononline database sources (i.e., listserv, personal communications, reference sections).

Eligibility Criteria

In order to be included in our analysis, a study needed to meet the following criteria: First, the study needed to have a true experimental design with random assignment to condition. Second, the study needed to compare fiction reading to either no reading or nonfiction reading. Third, the study needed to include at least one measure of social cognition, which, in line with one prior meta-analysis on this topic (Mumper & Gerrig, 2017), we operationalized as any measure testing the processes that underlie how one perceives, interprets, and responds to social information (Fiske & Taylor, 2013). The hypothesis that fiction reading improves social cognition has largely refrained from specifying which type of social cognition reading should improve. That is, researchers have asserted positive effects of fiction on multiple processes and skills that contribute to social behavior. Thus, here we include any study that measures social cognition, broadly defined in order to test the most general claim, that fiction has a positive causal effect on social cognition. This includes tests of mentalizing or theory-of-mind (i.e., mind reading, perspective-taking, cognitive empathy), experience sharing—sharing the internal affective experience of others (i.e., affective empathy, emotional contagion; Zaki & Ochsner, 2012), and social behavior (e.g., prosociality).

After an initial title and abstract screen, 52 full-texts were retrieved and assessed for eligibility, of which 14 studies ("study" being a published article or an unpublished dataset, which may contain multiple experiments with multiple effect sizes) were deemed eligible for inclusion (see Figure 1). The first author conducted the search and screen; the second author reviewed the eligibility assessment.

Data Extraction and Study Coding

We aimed to characterize not only the effect of fiction reading on social cognition, but also how study-level factors modulated this effect. For this reason, in addition to coding the necessary statistics to calculate effect sizes (n, M, SD), we coded studies for the following seven variables:

***Publication status* (published or unpublished).** Studies were considered published if they were published or in press in a peer-reviewed journal; dissertation studies were considered unpublished.

***Sample Type* (students, Amazon's Mechanical Turk Participants, or mixed sample).**

***Comparison group* (no reading or nonfiction).** Researchers have distinguished between and compared "literary" fiction to "popular" fiction (e.g., Kidd & Castano, 2013; Pino & Mazza, 2016) with the former described as focused on character development, and the latter on plot development (Kidd et al., 2016). That said, researchers have noted that the exact boundary between the two genres is unclear (Kidd et al., 2016). For this reason, we chose not to compare literary and popular fiction here. Furthermore, because the majority of studies we find use what would be characterized as "literary" as opposed to "popular" fiction, in studies that included both literary and popular fiction conditions, we included data only from the literary condition.

***Social-cognitive measure*.** See Table 1 for the measures used in each study.

***Dependent variable format* (performance-based or self-report).** Objective measures that assessed one's ability to accurately interpret

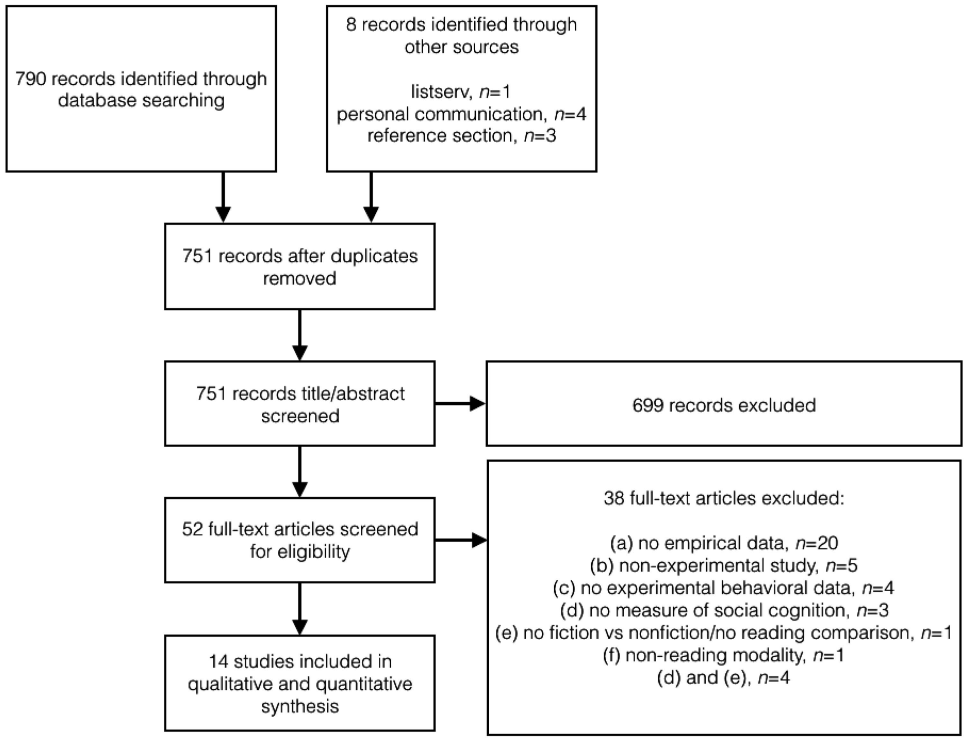

Figure 1. Flow diagram of studies through systematic review process.

information or behavior, or that directly measured social behavior (e.g., prosociality) were considered performance-based; subjective measures that asked participants to self-assess one's ability or tendency to engage in social–cognitive processes were considered self-report.

***Dependent variable social process* (mentalizing or experience sharing).** Mentalizing tasks assess the process by which we attribute and reason about the mental and emotional states of others (e.g., ToM, mind reading, perspective-taking, social perception, cognitive empathy), and experience sharing tasks assess one's ability or tendency to share the internal affective experience of others (e.g., affective empathy, emotional contagion; Zaki & Ochsner, 2012). Said otherwise, mentalizing tasks assess *understanding* or *consideration* of others' internal states while experience sharing tasks assess the ability or tendency to *vicariously experience* others' internal states. The one measure of pro-social behavior (Koopman, 2015) did not clearly fit in either of these categories so we did not include this effect size in the moderator analysis.

Participant characteristics. We coded the mean age of the entire sample and the percentage of female participants in the entire sample.

For unpublished data sets ($n = 12$ effect sizes), means and standard deviations were extracted using the authors' suggestions regarding data exclusion (e.g., dependent variable outliers; see Table 2). To avoid potential bias and to assess the effect's robustness to exclusion criteria, we also calculated effect sizes from unpublished data sets without excluding any data points (unless we received the data with outliers already removed), and reran all analyses. We report results using these data in the supplemental materials. Results remained consistent across both analyses.

Study coding was performed independently by the first and second author. Interrater reliability was calculated using Cohen's kappa for nominal data (e.g., publication status) and the intraclass correlation coefficient (ICC) for continuous data (e.g., M). Overall, reliability was high (Table S1). Disagreements in coding were resolved by discussion.

Statistical Analysis

We conducted all analyses in R (version 3.3.1; R Core Team, 2016) using the *metafor* package (Viechtbauer, 2010). All data and analysis code are made freely available on the Open Science Framework at osf.io/dz6qy.

Effect size calculation. Effect sizes were calculated as bias-corrected Hedges' g, representing the standardized mean difference between fiction reading and the comparison group such that positive effect sizes represent better performance in the fiction group. Scores that represented number of errors (e.g., DANVA2-AF) were reverse-scored to maintain this convention. We used raw unadjusted means and standard deviations, and compared only postreading between-groups scores. If not reported in the study, authors were contacted directly for this information.

Data synthesis. Summary effect size estimates represent inverse-variance weighted averages. Eleven out of the 14 studies we included in the meta-analysis report multiple studies with multiple effect sizes often from the same participants. This data structure, where effect sizes are derived from the same participants and nested within a larger study, violates the assumption of independence underlying traditional meta-analytic approaches (Lipsey

Table 1
Social Cognitive Measures

Measure	Abbreviation	Number ES (%)	Format	Domain	Description
Reading the Mind in the Eyes Task (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001)	RMET	17 (32.1)	Performance	Mentalizing	Identify mental state from image depicting only the eye region of actors.
Interpersonal Reactivity Index—Perspective-Taking (Davis, 1983)	IRI-PT	6 (11.3)	Self-Report	Mentalizing	Indicate tendency or ability to take others' perspectives.
Interpersonal Reactivity Index—Empathic Concern (Davis, 1983)	IRI-EC	5 (9.4)	Self-Report	Emotion Sharing	Indicate tendency to experience compassion and concern for others.
False-Belief Task (Converse, Lin, Keysar, & Epley, 2008; Rowe, Bullock, Polkey, & Morris, 2001)	FB	5 (9.4)	Performance	Mentalizing	Infer story character's belief or action based on that character's false-belief (e.g., of an object's location).
Empathy (Johnson, Jasper, Griffin, & Huffman, 2013; Koopman, 2015)	–	3 (5.7)	Self-Report	Emotion Sharing	Rate empathy felt towards a group of people (Arab-Muslims in Johnson, Cushman, Borden, & McCune, 2013; Depressed/bereaved individuals in Koopman, 2015).
Imposing Memory Task (Kinderman, Dunbar, & Bentall, 1998)	IMT	3 (5.7)	Performance	Mentalizing	Infer beliefs and emotions of story characters at increasing levels of complexity.
Diagnostic Analysis of Nonverbal Sensitivity 2 — Adults Faces Test (Nowicki & Duke, 1994)	DANVA2-AF	2 (3.8)	Performance	Mentalizing	Identify emotion from image of actors' entire face as expressing anger, fear, sadness, or happiness at differing levels of intensity.
Yoni Task (Shamay-Tsoory & Aharon-Peretz, 2007)		2 (3.8)	Performance	Mentalizing	Infer the mental states of a cartoon outline of a face based on gaze and facial expression.
Emotion Attribution Task (Blair & Cipolotti, 2000)	EAT	1 (1.9)	Performance	Mentalizing	Identify story character's emotion based on vignette of emotional situation.
Empathy Quotient—Cognitive Empathy (Baron-Cohen & Wheelwright, 2004)	EQ-CE	1 (1.9)	Self-Report	Mentalizing	Indicate ability to take others' perspectives.
Empathy Quotient—Emotional Reactivity (Baron-Cohen & Wheelwright, 2004)	EQ-ER	1 (1.9)	Self-Report	Emotion Sharing	Indicate tendency to have emotional response to others' mental state.
Faces Test (Baron-Cohen, Wheelwright, & Jolliffe, 1997)	FT	1 (1.9)	Performance	Mentalizing	Identify emotion from image of actress' face.
Multifaceted Empathy Test—Cognitive Empathy (Dziobek et al., 2008)	MET-CE	1 (1.9)	Performance	Mentalizing	Identify mental state or emotion from IAPs pictures.
Multifaceted Empathy Test—Explicit Emotional Empathy (Dziobek et al., 2008)	MET-EEE	1 (1.9)	Self-Report	Emotion Sharing	Indicate how strongly one feels for individuals in distress depicted in an image.
Multifaceted Empathy Test—Implicit Emotional Empathy (Dziobek et al., 2008)	MET-IEE	1 (1.9)	Self-Report	Emotion Sharing	Indicate one's emotional arousal in response to image of individuals in distress.
Prosocial Behavior (Koopman, 2015)	Prosociality	1 (1.9)	Performance	–	Participants given the opportunity to donate money received from participation to a charity.
Social-Reasoning Task (Mar, 2007)	SRT	1 (1.9)	Performance	Mentalizing	Infer mental state of story character involved in social situation.
Toronto Empathy Questionnaire (Koopman, 2015)	TEQ	1 (1.9)	Self-Report	Emotion Sharing	Indicate empathic response to others.

Note. ES = effect size.

& Wilson, 2001). Researchers have accounted for such dependencies in several ways; however, these methods can compromise the validity of meta-analytic estimates or result in the loss of information and power (see Cheung, 2014 for discussion). We employed two methods in order to incorporate all the effect sizes found in our search, and appropriately account for the two types of dependencies within the dataset (i.e., effect sizes nested within studies or "hierarchical effects," and effect sizes derived from overlapping participant samples or "correlated effects"; Tanner-Smith & Tipton, 2014). First, we implemented a multilevel random-effects meta-analytic model accounting for variance in the observed effect sizes (level 1), variance between effect sizes within a study (level 2), and variance between studies (level 3; Cheung, 2014; Konstantopoulos, 2011; Van den Noortgate, López-López, Marín-Martínez, & Sánchez-Meca, 2015). Unlike standard two-level meta-analyses in which random variance is estimated only for between-study differences, since each study is considered to contribute an independent effect, using three-level models to esti-

Table 2
Study Characteristics

Study	Number of ES	Sample	Fiction n^a	No Reading n^a	Nonfiction n^a	Fiction Material (dose)	Nonfiction Material (dose)	Measure(s)[b]	Data Exclusion	Publication Status
Bal & Veltkamp, 2013	2	Student	86	–	77	Section of *The Adventure of the Six Napoleons*, Arthur Conan Doyle; First chapter of *Blindness*, José Saramago (one text)	Articles from Dutch newspapers *De Volkskrant* and *NRC Handelsblad* (two to five texts)	IRI-EC	Inaccurate summary of reading	Published
Black & Barnes, 2015	1	Student	60	–	60	*The Runner*, Don Delillo; *Puppy*, George Saunders (two texts)	*The Story of the Most Common Bird in the World*, Smithsonian Magazine; *The CIA's Most Highly-Trained Spies Weren't Even Human*, Smithsonian Magazine (two texts)	RMET	Reading time outlier, RMET time outlier	Published
Djikic, Oatley, & Moldoveanu, 2013	3	Student	48	–	46	*The Echo*, Paul Bowles; *Night Club*, Katharine Brush; *My Oedipus Complex*, Frank O'Connor; *A Country Love Story*, Jean Stafford; *In the Zoo*, Jean Stafford; *Beyond a Glass Mountain*, Wallace Stegner; *The Wind and Snow of Winter*, Clark van Tilburg; *Prohibition*, Glenway Wescott (one text)	*Why Do We Laugh?*, Henri Bergson; *Science and Literature*, John Burroughs; *What Makes a Woman Beautiful?*, Havelock Ellis; *Dreams of the Death of Beloved Persons*, Sigmund Freud; *Castles in Spain*, John Galsworthy; *Nonmoral Nature*, Stephen Jay Gould; *Killing for Sport*, George Bernard Shaw; *East and West*, Rabindranath Tagore (one text)	IRI-EC, IRI-PT, RMET	Poor reading comprehension test performance	Published
Johnson et al., 2013	4	Student, MTurk	67	–	70	*Saffron Dreams*, Shila Abdullah (8-page excerpt)	"Brief history of the automobile" (one text)	Empathy, IRI-PT	None reported	Published

(*table continues*)

Table 2 (*continued*)

Study	Number of ES	Sample	Fiction n^a	No Reading n^a	Nonfiction n^a	Fiction Material (dose)	Nonfiction Material (dose)	Measure(s)[b]	Data Exclusion	Publication Status
Kidd & Castano, 2013	6	MTurk	195	174	43	*The Runner*, Don DeLillo; *Blind Date*, Lydia Davis; *Chameleon*, Anton Chekhov; *The Round House*, Louise Erdrich; *The Tiger's Wife*, Téa Obreht; *Salvage the Bones*, Jesmyn Ward; *Corrie*, Alice Munroe; *Leak*, Sam Ruddick; *Nothing Living Lives Alone*, Wendell Berry; *Uncle Rock*, Dagoberto Gilb; *The Vandercook*, Alice Mattinson (one text)	*How the Potato Changed the World*, Charles C. Mann; *Bamboo Steps Up*, Cathie Gandel; *The Story of the Most Common Bird in the World*, Rob Dunn (one text)	DANVA2-AF, FB, RMET, Yoni Task	Reading time outlier, ART-foil outlier, dependent measure outlier	Published
Kidd, Ongis, & Castano, 2016	1	MTurk	109	115	–	*Nothing Living Lives Alone*, Wendell Berry; *Chameleon*, Anton Chekhov; *Pacific*, Tom Drury; *Uncle Rock*, Dagoberto, Gilb; *The End of the Point*, Elizabeth Graver; *The Vandercook*, Alice Mattison; *The Good Lord Bird*, James McBride; *Someone*, Alice McDermott; *Corrie*, Alice Munro; *Bleeding Edge*, Thomas Pynchon; *Leak*, Sam Ruddick (one full text or text excerpt)	–	RMET	Reading time outlier, ART-foil outlier, prior participation in similar study, Non-native English speakers, self-reported distraction/technical difficulty	Published
Koopman, 2015	3	Student	86	–	40	*Counterpoint*, Anna Enquist; *Over de Liefde [On Love]*, Doeschka Meijsing (two text excerpts)	*Malignant Sadness*, Lewis Wolpert; *The Bereaved Parents' Survival Guide*, Juliet Cassuto Rothman (two text excerpts)	Empathy, Prosociality, TEQ	Dependent measure outlier, five or more missing variables, incorrect responses to recall items	Published
Liu & Want, 2015	3	Student	51	–	52	*Chameleon*, Anton Chekov (one text)	*Bamboo Steps Up*, Cathie Gandel (one text)	IRI-EC, IRI-PT, RMET	Dependent measure outlier, reading time outlier	Unpublished

Table 2 (*continued*)

Study	Number of ES	Sample	Fiction n^a	No Reading n^a	Nonfiction n^a	Fiction Material (dose)	Nonfiction Material (dose)	Measure(s)[b]	Data Exclusion	Publication Status
Mar, 2007	1	Student	67	–	77	*The Orlov-Sokolovs*, Ludmila Ulitskaya (one text)	*Getting In: The Social Logic of Ivy League Admission*, Malcolm Gladwell (one text)	SRT	< 9 years English fluency, dependent measure outlier	Unpublished Dissertation
Mar	2	Student	64	–	62	Same as above	Same as above	IRI-EC, IRI-PT	ART-foil outlier, dependent measure outlier	Unpublished
Panero et al., 2016[c]	10	MTurk and Student	357	229	207	Kidd and Castano (2013) readings; *Puppy*, George Saunders (one text)	Kidd and Castano (2013) readings; *The CIA's Most Highly-Trained Spies Weren't Even Human*, Smithsonian Magazine; excerpt of *The Kid*, Dan Savage; *M*, John Sack; *Frank Sinatra Has a Cold,* Gay Talese (one text)	DANVA2-AF, FB, IMT, RMET, Yoni Task	Sample restricted to United States, not completing the study, reading time outlier, dependent measure outlier, ART-foil outlier (one research group also used Poor performance on reading material memory test and reading manipulation check failure)	Published/Unpublished
Pino & Mazza, 2016	8	Student	74	–	67	*Tenth of December*, George Saunders; *Io e Te*, Niccolò Ammaniti (one text)	*Intervista ad Oriana Fallaci*, Oriana Fallaci; *Wave*, Sonali Deraniyagala (one text)	EAT, EQ-CE, EQ-ER, FB, FT, MET-CE, MET-EEE, MET-IEE	Poor reading comprehension test performance	Published

(*table continues*)

Table 2 (*continued*)

Study	Number of ES	Sample	Fiction n^a	No Reading n^a	Nonfiction n^a	Fiction Material (dose)	Nonfiction Material (dose)	Measure(s)[b]	Data Exclusion	Publication Status
Samur, Tops, & Koole, 2017	6	MTurk	301	188	304	*The Runner*, Don DeLillo; *Blind Date*, Lydia Davis; *Chameleon*, Anton Chekov; *The Tiger's Wife*, Téa Obreht; *Uncle Rock*, Dagoberto Gilb; *The Vandercook*, Alica Mattison; *Corrie*, Alice Munroe; *Leak*, Sam Ruddick; *Nothing Living Lives Alone*, Wendell Berry (one text)	*How the Potato Changed the World*, Charles C. Mann; *Bamboo Steps Up*, Cathie Gandel; *The Story of the Most Common Bird in the World*, Rob Dunn; *How the Chicken Conquered the World*, Jerry Adler, Andrew Lawler; *Mistletoe: The Evolution of a Christmas Tradition*, Rob Dunn; *The Venus Flytrap's Lethal Allure*, Abigail Tucker; *Exploring the Titanic of the Ancient World*, Jo Marchant; *Can the Siberian Tiger Make a Comeback*, Matthew Shaer (one text)	RMET	Reading time outlier, ART-foil outlier, dependent measure outlier	Published
Weisberg	3	Student	50	32	–	*Brighton Rock*, Graham Greene; Empire *of the Sun*, J. G. Ballard; *Remains of the Day*, Kazuo Ishiguro (one text excerpt)	–	IMT, IRI-PT, RMET	Poor performance on reading material memory test, ART-foil outlier, dependent measure outlier	Unpublished

[a] Number represents the sum of independent participants across all effect sizes in a given study. [b] See Table 1 for measure names and descriptions. [c] Eight of the effect sizes from Panero et al. (2016) are not reported in the main manuscript, and thus were considered unpublished effect sizes in the moderator analysis.

mate random variation at level 2 accounts for dependence (i.e., clustering of effect sizes) among effect sizes from the same study (Konstantopoulos, 2011; Scammacca, Roberts, & Stuebing, 2014). Though this method accounts for the hierarchical dependence among effect sizes nested within studies, it assumes independent sampling errors within data clusters, which is violated in our dataset through overlapping samples in effect size estimates (i.e., fiction participants are compared to both no reading and nonfiction participants, and between-groups comparisons are made on multiple social–cognitive measures; Tanner-Smith, Tipton, & Polanin, 2016). To account for these correlated effects, we generated cluster-robust standard errors, statistical tests, and confidence intervals on estimates from the three-level meta-analytic model (Cameron & Miller, 2015; Hedges, Tipton, & Johnson, 2010).

The presence of variability among effect sizes may speak to how study-level factors (e.g., comparison group, dependent measure) influence whether and how fiction reading impacts social cognition. Thus, we assessed the presence of heterogeneity among the effect sizes with the Q statistic (Cochran, 1954). Statistically significant Q values suggest that the effect sizes are not estimating a common population mean; that is, the effect sizes differ from each other beyond what would be expected from sampling error, for example, because of significant moderating factors across the studies. In this case, the use of a random-effects (vs. fixed-effects) model is appropriate, as well as testing for moderating factors that may be contributing to the heterogeneity between effect sizes. Finally, we report total I^2 and its components: $I^2_{\text{Level 2}}$ and $I^2_{\text{Level 3}}$. I^2 represents the percentage of variability that is due to true heterogeneity between effect sizes rather than sampling error, with $I^2_{\text{Level 2}}$ and $I^2_{\text{Level 3}}$ representing within- and between-study heterogeneity, respectively. Large I^2 values indicate that a large proportion of variation between effect sizes is not due to chance, and instead that variance in effect sizes may be caused by systematic differences in study- or experiment-level factors, which we investigate with moderator analyses. Low I^2 values indicate that the variance among effect sizes are due to sampling variability (i.e., chance), and not because of systematic differences between studies. We use the benchmarks provided by Higgins and Green (2011) to interpret the magnitude of heterogeneity. Model parameters were estimated using restricted maximum likelihood (REML) estimation (Cheung, 2014; Viechtbauer, 2010).

Moderator analysis. We can determine if there is large, and meaningful variance between effect sizes by calculating the Q statistic and I^2 values. If so, we can then test whether variability among effect sizes may be accounted for by systematic differences within and between studies, by assessing how specific methodological or psychological factors influence the effect of fiction reading on social cognition. We investigated potential sources of heterogeneity by conducting moderator analyses with characteristics that varied between-studies (e.g., publication status) and within-studies (e.g., dependent variable format). Some study characteristics, for example, measure format, varied both within- and between-studies, with some studies including both performance-based and self-report measures, and other studies including only one or the other. In this scenario, because estimates may be confounded by between-study characteristics, we conducted follow-up analyses, reestimating the model using only those studies that contained both levels of a characteristic.

Sensitivity analyses. Effect sizes that markedly deviate from others (i.e., outliers), and greatly impact statistical model coefficients (i.e., influential) may distort summary statistics and lead to spurious conclusions. For this reason, we examined the data for influential outliers (Aguinis, Gottfredson, & Joo, 2013; Habeck & Schultz, 2015), which we defined as effect sizes with standardized residual values exceeding 3.0 (Cohen & Cohen, 2003) whose Cook's Distance values exceeded .078, which was determined using the formula, $4/(n\text{-}k\text{-}1)$ (Fox, 1991). We note that this is a conservative estimate of influence as others have suggested using a Cook's Distance value of >1 as a cutoff (Cohen & Cohen, 2003). We reran the multilevel model after excluding these effect sizes from the analysis, and report these findings alongside the model including these effect sizes. We conducted two other sensitivity analyses using a leave-one-out procedure to gauge the impact of each individual effect size (i.e., rerunning the multilevel model leaving out one effect size at a time) and study (i.e., rerunning the multilevel model leaving one study out at a time) on the overall effect and amount of heterogeneity.

Publication bias. We aimed to include all relevant data by casting a wide search through posts on listservs and direct communication with authors. As such, the final analyses include both published and unpublished data. We evaluated the possibility of publication bias in several ways. First, as described above, we tested whether publication status moderated the effect; that is, whether the effect sizes from published studies differed systematically from the effect sizes of unpublished studies.

Second, we inspected funnel plots, which depict the relation between effect sizes and the precision of the effects (i.e., their standard error). The rationale behind funnel plots is that larger, more precise studies at the top of the plot should scatter tightly around the mean or true effect size, while smaller, less precise studies at the bottom of the plot should scatter widely around the mean, creating a funnel shape. If publication limited our access to nonsignificant findings, then significant findings would be overrepresented in the analyses, whereas nonsignificant findings would be omitted from the analysis. The funnel plot allows us to evaluate this possibility, which would be observed as a lack of data points to the left of the mean and an overrepresentation of low-powered significant findings to the right of the mean. A disturbance of symmetry in this direction would be evidence for small-study effects—of which publication bias may be one source—that may artificially inflate our estimated effect size. While informative, we note several caveats related to visually inspecting funnel plots, including the subjectivity inherent in their interpretation, and the fact that they do not take into account the data's multilevel structure, which may lead to clusters of data points and asymmetry that could be misinterpreted as bias (Lau, Ioannidis, Terrin, Schmid, & Olkin, 2006). To formally evaluate funnel plot asymmetry, we performed Egger's regression test by including the standard error of the effect sizes estimates as a moderator in the multilevel models. That is, we evaluated whether the precision of the effect (i.e., the standard error) was associated with the magnitude of the effect (i.e., the effect size). Statistically significant standard error coefficients would suggest that effect sizes from high precision studies systematically differ from effect sizes from low precision studies, meaning that bias may be present.

Results

Study Characteristics

We obtained 53 effect sizes from 14 studies (see Table 2). Effect sizes were derived from 1,615 fiction participants and 1,843 control participants, 63.3% of which were female (SD = 14.0%, study sample range = 33.0–85.4%), with a mean age of 27.9 (SD = 6.8, study sample range = 18.9–37.5). Most effect sizes were derived from published studies (67.9% vs. 32.1% unpublished), using a student sample (54.7% vs. 34.0% from MTurk vs. 11.3% from a mixed sample), nonfiction comparison group (71.7% vs. 28.3% from a no reading comparison), performance-based measures (64.2% vs. 35.8% from self-report), and were from tasks assessing mentalizing (75.5% vs. 24.5% for experience sharing) (see Table 2). The most widely used social–cognitive measure was RMET, which comprised 32.1% of effect sizes. The distribution of the 53 effect sizes was positively skewed such that the distribution's tail extended rightward with several larger effect sizes favoring the hypothesis that fiction reading improves social cognition. Individual effect sizes ranged from −.55 (Bal & Veltkamp, 2013) to 1.51 (Johnson, Jasper, Griffin, & Huffman, 2013).

Meta-Analysis

Our primary question concerned whether fiction readers outperformed nonfiction/no readers across the 14 studies. Meta-analysis of the 53 effect sizes demonstrated that fiction readers outperformed nonfiction and no readers on social–cognitive tasks. This effect was small and statistically significant, g = .15, 95% CI [.02, .29], p = .029 (Table 3, Figure 2). Reestimating the model using effect sizes derived from the unpublished data sets with no data exclusions applied yielded the same findings in this analysis, g = .16, 95% CI [.03, .30], p = .021, and all subsequent analyses (see supplemental material). The Q statistic was significant, $Q(52)$ = 164.24, $p < .0001$, indicating the presence of heterogeneity. Total I^2 was 71% indicating a substantial amount of true variance (vs. sampling error) in effect size estimates, the majority of which came from within-study variance, $I^2_{\text{Level 2}}$ = 47%, versus between-study variance, $I^2_{\text{Level 2}}$ = 24%. Together, these results suggest that fiction reading marginally improves social cognition, and that there appear to be systematic differences among the effect sizes due in large part to factors that vary within-study (e.g., comparison group, dependent measure format, dependent measure process assessed).

The size of the effect was reliable across several robustness checks. Outlier and influence diagnostics identified one influential outlier (g = 1.51; Johnson, Jasper, et al., 2013). We reran the analysis with this effect size removed, which decreased the size of the overall effect, but did not change the significance of the findings, g = .13, 95% CI [.01, .26], p = .039. Leave-one-out analysis at the individual effect size level demonstrated that magnitude of the effect was robust to any one effect size (range g = .13–.17); however, the removal of the SRT from Mar (2007) reduced the findings to a trend level of significance (p = .051). True effect size variance remained substantial (low I^2 = 65% [$I^2_{\text{Level 2}}$ = 41%, $I^2_{\text{Level 3}}$ = 24%]; high I^2 = 72% [$I^2_{\text{Level 2}}$ = 35%, $I^2_{\text{Level 3}}$ = 37%]. Similarly, leave-one-out analysis at the study level demonstrated that the magnitude of the effect was robust to any one study (range g = .12–.19). However, removal of each of the following four studies reduced the findings to a trend level of significance: Johnson, Cushman, Borden, & McCune, (2013) (p = .053), Kidd and Castano (2013) (p = .053), Pino and Mazza (2016) (p = .063), or Mar (2007) (p = .051; this paper reported only one effect size so this finding is the same as reported above in the leave-one-effect-size-out analysis). True effect size variance remained substantial (low I^2 = 63% [$I^2_{\text{Level 2}}$ = 39%, $I^2_{\text{Level 3}}$ = 24%]; high I^2 = 73% [$I^2_{\text{Level 2}}$ = 48%, $I^2_{\text{Level 3}}$ = 25%]). Analysis of the effect sizes derived from the unpublished data sets with no data exclusions applied was more robust to any one effect size or study (supplemental material).

Table 3
Meta-Analytic Results

Variable	Number of Studies	Number of ES	g	95% CI	SE	t	p	Q
Overall Estimate	14	53	.15*	.02, .29	.06	2.46	.029	164.24***
Publication Status						1.29	.223	154.60***
Published	10	36	.19*	.04, .34	.07			
Unpublished	5	17	.08	−.10, .25	.08			
Sample[a]						.27	.796	141.57***
Mechanical Turk	5	18	.15*	.02, .27	.06			
Student	11	29	.17+	−.02, .36	.09			
Comparison						1.05	.314	164.23***
No Reading	5	15	.21**	.09, .32	.05			
Nonfiction	12	38	.13	−.04, .31	.08			
Dependent Variable Format						1.19	.259	164.21***
Performance	10	34	.21*	.01, .40	.09			
Self-Report	8	19	.08	−.10, .25	.08			
Dependent Variable Process						.33	.750	161.90***
Emotion Sharing	7	12	.20	−.19, .59	.18			
Mentalizing	12	40	.13	−.05, .31	.08			
Participant Characteristics								
Age	14	53	b = −.003	−.017, .011	.007	.47	.649	158.01***
Percent Female	14	53	b = .0004	−.007, .008	.003	.14	.895	161.15***

Note. ES = effect sizes. Number of Studies within a variable may exceed N = 14 as some studies contained both levels of the variable.
[a] Six effect sizes from Panero et al. (2016) contained mixed samples with student and Mechanical Turk participants and were not included in the moderator analysis.
+ $p \leq .10$. * $p \leq .05$. ** $p \leq .01$. *** $p \leq .001$.

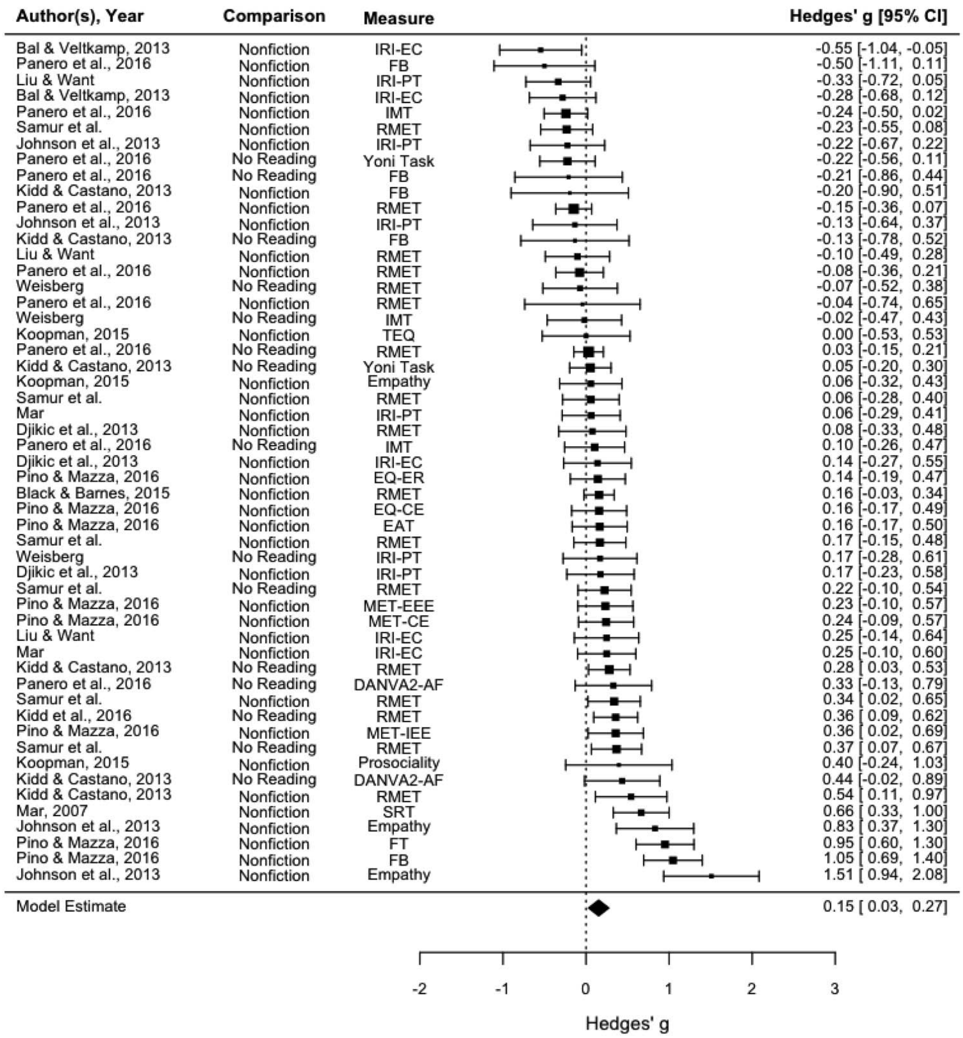

Figure 2. Forest plot of the results.

Given the large number of effect sizes derived from RMET, and the recent controversy involving the replicability of findings with this task, we evaluated the effect of fiction reading specifically on RMET performance (number of studies = 8, number of effect sizes = 17). We observed that fiction reading improves RMET performance; this was a small effect, and statistically significant at a trend level, $g = .13$, 95% CI [–.02, .28], $p = .084$, with moderate heterogeneity, $Q(16) = 29.28$, $p = .022$, $I^2 = 48\%$ [$I^2_{\text{Level 2}} = 14\%$, $I^2_{\text{Level 3}} = 34\%$].

Moderator Analysis

To investigate potential sources of heterogeneity, we evaluated the effect of the following factors on the meta-analytic estimate: sample type (students vs. Amazon's Mechanical Turk participants), comparison group (no reading vs. nonfiction), measure format (performance-based vs. self-report), measure process (mentalizing vs. experience sharing), and participant characteristics, including age and percentage of female participants. None of these factors moderated the effect (see Table 3). Findings remained the same when excluding the influential outlier. To exclude the possibility that within-study moderator analyses (i.e., comparison group, measure format, measure process) were influenced by between-study differences, we reran these moderator analyses including only those studies that included both levels of these moderators. Findings remained the same. Together, these results suggest that though there exists substantial inconsistency among the effect sizes, none of the factors investigated here accounted for the heterogeneity.

Publication Bias

We found no evidence of publication bias. Although published effect sizes yielded larger effects, $g = .19$, 95% CI [.04, .34], than unpublished ones, $g = .08$, 95% CI [−.10, .25], publication status did not moderate the effect of fiction reading ($p = .223$) (see Table 3). Visual inspection of the funnel plot revealed some asymmetry. However, this asymmetry is the *opposite* of what would be expected from publication bias: there were *more* data points from less precise studies to the left of the

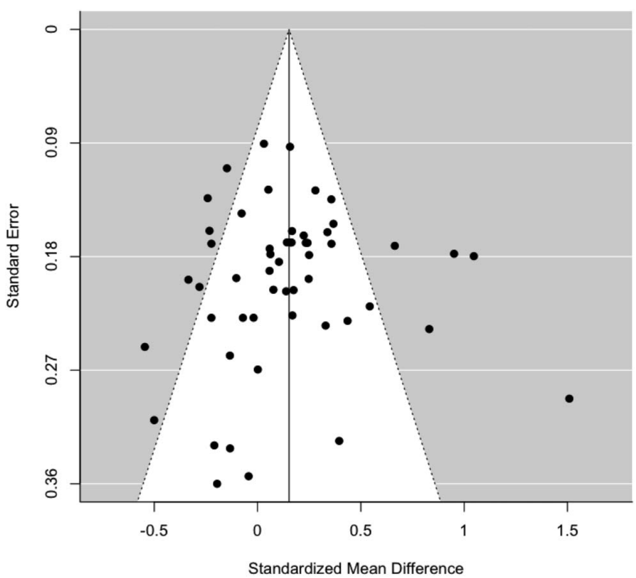

Figure 3. Funnel plot of the results.

mean effect (see Figure 3). That is, smaller, less precise studies, tended to yield smaller effects, or effects that suggest fiction reading had no effect on or impaired social cognition. In the case of publication bias, we would expect to see a greater presence of smaller studies yielding large effects (i.e., to the right of the mean effect on the plot). Consistent with this interpretation, the slope of Egger's regression test for funnel plot asymmetry was negative, but not significant, $b = -.20$, $SE = .60$, $p = .739$, meaning that the precision of the measured effect was not significantly related to the magnitude of the effect.

Discussion

Does fiction reading improve social cognition? The current meta-analysis of experimental studies suggests that it does. We find that fiction reading leads to a small ($g = .15–.16$), but statistically significant improvement in social–cognitive performance compared to nonfiction reading or no reading. These findings support a causal view of fiction's effect on social–cognitive ability; that is, fiction reading is correlated with social–cognitive ability (Mumper & Gerrig, 2017) because fiction reading causally improves social cognition. These effects were robust across all sensitivity analyses and did not appear to be the result of publication bias, suggesting that the impact of reading on social cognition may be small, but reliable.

The small size of the effect raises the question of how meaningful it is. We argue that this effect has the potential to be very meaningful. The magnitude of an effect does not determine its practical impact (Cooper, 2008). Social–cognitive skills have been shown to positively impact social connection across the life span (e.g., Goldstein, Vezich, & Shapiro, 2014; Slaughter, Imuta, Peterson, & Henry, 2015), particularly in clinical populations (e.g., Fett et al., 2011). Strong social connections can significantly improve well-being (Reis, Sheldon, Gable, Roscoe, & Ryan, 2000; Ryan & Deci, 2000), stave off physical illness (Yang et al., 2016), and enhance longevity (Holt-Lunstad, Smith, & Layton, 2010).[1] Thus, any method that enhances social–cognitive skills in the general population or in individuals with social–cognitive deficits is worthwhile and deserving of additional research—especially when this method is cost-effective, easily disseminated, and well-tolerated.

It is also important to consider the possibility that fiction may have an even larger impact with more immersive, longitudinal reading experiences. Almost all of the studies in this meta-analysis required participants to read only one short fiction story. Longer periods of reading may yield larger or longer-lasting effects; indeed, the one study included here that required participants to read an entire book yielded some of the largest effect sizes (Pino & Mazza, 2016). As is, it is not clear whether improvements in social cognition from fiction reading represent a short-term change, along the lines of a priming effect, versus an enduring change in social–cognitive ability. Other than Bal and Veltkamp (2013) who found positive changes in empathy one week after reading, few studies have assessed the durability of improvements.

Of course, the impact of this method hinges on the extent to which the effects transfer. Do improvements in social cognition generalize to improved real-world or day-to-day social behavior? Evidence that fiction reading increases pro-social behavior, and not just social cognition, suggests it may (Johnson, 2012; Johnson, Cushman et al., 2013; Koopman, 2015), but further work is needed.

The current findings help to resolve the debate over *whether* fiction reading improves social cognition. However, we still do not know *how* fiction reading improves social cognition, and what factors may influence this association. Mar (2015) has proposed two routes (also see Oatley, 2016). In the *process* route, readers get to practice and strengthen their social–cognitive skills because reading involves repeated simulation of social–cognitive processes (Mar & Oatley, 2008; Oatley, 2016; Tamir et al., 2016); in the *content* route, reading provides concrete knowledge about human psychology and social interaction. Future work should elucidate the independent and joint contribution of these mechanisms.

None of the moderators tested in these analyses offer any new insight into this question. That is, the effect of fiction on social cognition did not statistically differ between the different levels of any given moderating variable. However, this meta-analysis, while sufficiently powered to detect a main effect of fiction reading on social cognition, is likely underpowered to fully assess moderating factors. Power is often hampered in moderator analyses as the size of moderator effects are usually smaller than main effects, and the sample sizes within the groups being compared are smaller than the total sample size (Borenstein, Hedges, Higgins, & Rothstein, 2011). Further, meta-analytic simulations have shown that when the proportion of the moderator is not equal among effect sizes (e.g., 28% no reading comparison vs. 72% nonfiction comparison) and heterogeneity is high, power is substantially compromised even in the case of a strong moderator effect (Hempel et al., 2013). Thus, the null moderator findings should not be taken as strong evidence *against* a lack of effect from the variables we examine.

[1] We note that reading books has independently been shown to contribute to similar outcomes, namely, reduced risk of mortality (Bavishi, Slade, & Levy, 2016).

Our results can thus only offer a *suggestion* as to which study factors may influence the effect of fiction. For example, the effect of fiction on social cognition was larger when compared to no reading versus nonfiction reading. Indeed, *if* fiction's causal impact depends on the extent to which a text provokes readers to consider mental states (Kotovych, Dixon, Bortolussi, & Holden, 2011; Peskin & Astington, 2004), then many forms of nonfiction (e.g., memoir) *may* likewise improve social cognition. This may account for the attenuated effect of fiction when compared to nonfiction, and offers additional reason to assess the features of fiction that allow it to improve social cognition.

Individual difference factors may also moderate the causal relation between fiction and social cognition. Given the same text, some readers may be more likely to benefit from fiction than others. Reading is an active experience, requiring willful participation by the reader (Gerrig & Wenzel, 2015). Thus, the benefits to social cognition may depend on the quality of a reader's engagement with a text and motivation to understand the characters (Keen, 2006). For example, fiction's impact may depend on a reader's propensity to be transported into narratives, generate imagery while reading, or to simulate other minds (Bal & Veltkamp, 2013; Calarco, Fong, Rain, & Mar, in press; Johnson, 2012; Johnson, Cushman, et al., 2013; Tamir et al., 2016). In the absence of this type of reader engagement, fiction is unlikely to effect any change at all. Furthermore, one's existing knowledge base, expertise, or age of exposure may determine how likely one is to benefit from fiction reading (e.g., Stanovich, 1986). If so, prior social–cognitive ability would also moderate fiction's impact. While we were not able to test these factors here, we recommend that future studies measure the role that individual differences play in moderating the effect of fiction reading on social cognition.

While we show here that fiction effects a small causal improvement of social cognition, it is also likely the reverse causal relation exists. That is, fiction reading and social cognition might form a mutually facilitating and reinforcing pathway, akin to a "Matthew Effect" (Merton, 1968; Stanovich, 1986; Walberg & Tsai, 1983). Socially skilled individuals may gravitate toward fiction due to its social content more than less-skilled individuals (see Barnes, 2012). In doing so, readers further differentiate their social–cognitive skills from nonreaders as part of a self-reinforcing cycle.

In summary, we find that fiction reading leads to a small improvement in social cognition. However, additional work is needed to tease apart the nature of fiction's effect on social cognition. Our findings offer a benchmark for conducting adequately powered work in this domain.[2] We recommend that in future work, researchers design studies with more robust reading manipulations, investigate fiction's impact on real-world or day-to-day social functioning, and focus on the causal mechanisms underlying this effect. Ultimately, a better understanding of why fiction reading improves social cognition will provide researchers a way of boosting the magnitude of the effect, which in turn, will provide clinicians, educators, policymakers, and parents, a way of maximizing its potential impact.

[2] Assuming an effect size of .15, $\alpha = .05$, one-tailed test, with an independent two-sample design (as is true of many of the studies reviewed here), $N = 1102$ (i.e., $n = 551$ per group) would be needed to achieve a power level of .80.

References

References marked with an asterisk (*) indicate studies included in the meta-analysis.

Aguinis, H., Gottfredson, R. K., & Joo, H. (2013). Best-Practice Recommendations for Defining, Identifying, and Handling Outliers. *Organizational Research Methods, 16,* 270–301. http://dx.doi.org/10.1177/1094428112470848

*Bal, P. M., & Veltkamp, M. (2013). How does fiction reading influence empathy? An experimental investigation on the role of emotional transportation. *PLoS ONE, 8*(1), e55341. http://dx.doi.org/10.1371/journal.pone.0055341

Barnes, J. L. (2012). Fiction, imagination, and social cognition: Insights from autism. *Poetics, 40,* 299–316. http://dx.doi.org/10.1016/j.poetic.2012.05.001

Baron-Cohen, S., Wheelwright, S., & Jolliffe, A. T. (1997). Is there a "language of the eyes"? Evidence from normal adults, and adults with autism or Asperger syndrome. *Visual Cognition, 4,* 311–331. http://dx.doi.org/10.1080/713756761

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: An investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders, 34,* 163–175. http://dx.doi.org/10.1023/B:JADD.0000022607.19833.00

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, 42,* 241–251. http://dx.doi.org/10.1111/1469-7610.00715

Bavishi, A., Slade, M. D., & Levy, B. R. (2016). A chapter a day: Association of book reading with longevity. *Social Science & Medicine, 164,* 44–48. http://dx.doi.org/10.1016/j.socscimed.2016.07.014

*Black, J. E., & Barnes, J. L. (2015). The effects of reading material on social and non-social cognition. *Poetics, 52,* 32–43. http://dx.doi.org/10.1016/j.poetic.2015.07.001

Blair, R. J. R., & Cipolotti, L. (2000). Impaired social response reversal: A case of "acquired sociopathy." *Brain: A Journal of Neurology, 123,* 1122–1141. http://dx.doi.org/10.1093/brain/123.6.1122

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. (2011). *Introduction to Meta-Analysis*. Somerset, NJ: Wiley. Retrieved from http://public.eblib.com/choice/publicfullrecord.aspx?p=4041330

Calarco, N., Fong, K., Rain, M., & Mar, R. (in press). Absorption in narrative fiction and its possible impact on social abilities. In F. Hakemulder, M. M. Kuijpers, E. S. H. Tan, K. Balint, & M. M. Doicaru (Eds.), *The handbook of narrative absorption*. Amsterdam, the Netherlands: John Benjamins. http://dx.doi.org/10.1075/lal.27.15cal

Cameron, A. C., & Miller, D. L. (2015). A Practitioner's Guide to Cluster-Robust Inference. *The Journal of Human Resources, 50,* 317–372. http://dx.doi.org/10.3368/jhr.50.2.317

Cheung, M. W.-L. (2014). Modeling dependent effect sizes with three-level meta-analyses: A structural equation modeling approach. *Psychological Methods, 19,* 211–229. http://dx.doi.org/10.1037/a0032968

Cochran, W. G. (1954). The combination of estimates from different experiments. *Biometrics, 10,* 101–129.

Converse, B. A., Lin, S., Keysar, B., & Epley, N. (2008). In the mood to get over yourself: Mood affects theory-of-mind use. *Emotion, 8,* 725–730. http://dx.doi.org/10.1037/a0013283

Cohen, J., & Cohen, J. (Eds.). (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: L. Erlbaum Associates.

Cooper, H. (2008). The Search for Meaningful Ways to Express the Effects of Interventions. *Child Development Perspectives, 2,* 181–186. http://dx.doi.org/10.1111/j.1750-8606.2008.00063.x

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology, 44,* 113–126. http://dx.doi.org/10.1037/0022-3514.44.1.113

*Djikic, M., Oatley, K., & Moldoveanu, M. C. (2013). Reading other minds: Effects of literature on empathy. *Scientific Study of Literature, 3,* 28–47. http://dx.doi.org/10.1075/ssol.3.1.06dji

Dziobek, I., Rogers, K., Fleck, S., Bahnemann, M., Heekeren, H. R., Wolf, O. T., & Convit, A. (2008). Dissociation of cognitive and emotional empathy in adults with Asperger syndrome using the Multifaceted Empathy Test (MET). *Journal of Autism and Developmental Disorders, 38,* 464–473. http://dx.doi.org/10.1007/s10803-007-0486-x

Fett, A.-K. J., Viechtbauer, W., Dominguez, M. D., Penn, D. L., van Os, J., & Krabbendam, L. (2011). The relationship between neurocognition and social cognition with functional outcomes in schizophrenia: A meta-analysis. *Neuroscience and Biobehavioral Reviews, 35,* 573–588. http://dx.doi.org/10.1016/j.neubiorev.2010.07.001

Fiske, S. T., & Taylor, S. E. (2013). *Social cognition: From brains to culture* (2nd ed.). Los Angeles, California: SAGE. http://dx.doi.org/10.4135/9781446286395

Fox, J. (1991). *Regression diagnostics*. Newbury Park, California: Sage. http://dx.doi.org/10.4135/9781412985604

Gerrig, R. J., & Wenzel, W. G. (2015). The role of inferences in narrative experiences. In E. J. OBrien, A. E. Cook, & J. Lorch (Eds.), *Inferences during reading* (pp. 362–385). Cambridge, England: Cambridge University Press. http://dx.doi.org/10.1017/CBO9781107279186.017

Goldstein, N. J., Vezich, I. S., & Shapiro, J. R. (2014). Perceived perspective taking: When others walk in our shoes. *Journal of Personality and Social Psychology, 106,* 941–960. http://dx.doi.org/10.1037/a0036395

Habeck, C. W., & Schultz, A. K. (2015). Community-level impacts of white-tailed deer on understorey plants in North American forests: A meta-analysis. *AoB Plants, 7,* plv119. http://dx.doi.org/10.1093/aobpla/plv119

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods, 1,* 39–65. http://dx.doi.org/10.1002/jrsm.5

Hempel, S., Miles, J. N., Booth, M. J., Wang, Z., Morton, S. C., & Shekelle, P. G. (2013). Risk of bias: A simulation study of power to detect study-level moderator effects in meta-analysis. *Systematic Reviews, 2,* 107. http://dx.doi.org/10.1186/2046-4053-2-107

Higgins, J. D., & Green, S. (2011). *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0* [updated March 2011]. The Cochrane Collaboration. Retrieved from http://handbook.cochrane.org

Holt-Lunstad, J., Smith, T. B., & Layton, J. B. (2010). Social relationships and mortality risk: A meta-analytic review. *PLoS Medicine, 7*(7), e1000316. http://dx.doi.org/10.1371/journal.pmed.1000316

Johnson, D. R. (2012). Transportation into a story increases empathy, prosocial behavior, and perceptual bias toward fearful expressions. *Personality and Individual Differences, 52,* 150–155. http://dx.doi.org/10.1016/j.paid.2011.10.005

Johnson, D. R., Cushman, G. K., Borden, L. A., & McCune, M. S. (2013). Potentiating empathic growth: Generating imagery while reading fiction increases empathy and prosocial behavior. *Psychology of Aesthetics, Creativity, and the Arts, 7,* 306–312. http://dx.doi.org/10.1037/a0033261

*Johnson, D. R., Jasper, D. M., Griffin, S., & Huffman, B. L. (2013). Reading Narrative Fiction Reduces Arab-Muslim Prejudice and Offers a Safe Haven From Intergroup Anxiety. *Social Cognition, 31,* 578–598. http://dx.doi.org/10.1521/soco.2013.31.5.578

Keen, S. (2006). A Theory of Narrative Empathy. *Narrative, 14,* 207–236. http://dx.doi.org/10.1353/nar.2006.0015

*Kidd, D. C., & Castano, E. (2013). Reading literary fiction improves theory of mind. *Science, 342,* 377–380. http://dx.doi.org/10.1126/science.1239918

*Kidd, D., Ongis, M., & Castano, E. (2016). On literary fiction and its effects on theory of mind. *Scientific Study of Literature, 6,* 42–58. http://dx.doi.org/10.1075/ssol.6.1.04kid

Kinderman, P., Dunbar, R., & Bentall, R. P. (1998). Theory-of-mind deficits and causal attributions. *British Journal of Psychology, 89,* 191–204. http://dx.doi.org/10.1111/j.2044-8295.1998.tb02680.x

Konstantopoulos, S. (2011). Fixed effects and variance components estimation in three-level meta-analysis. *Research Synthesis Methods, 2,* 61–76. http://dx.doi.org/10.1002/jrsm.35

*Koopman, E. M. (2015). Empathic reactions after reading: The role of genre, personal factors and affective responses. *Poetics, 50,* 62–79. http://dx.doi.org/10.1016/j.poetic.2015.02.008

Kotovych, M., Dixon, P., Bortolussi, M., & Holden, M. (2011). Textual determinants of a component of literary identification. *Scientific Study of Literature, 1,* 260–291. http://dx.doi.org/10.1075/ssol.1.2.05kot

Lau, J., Ioannidis, J. P. A., Terrin, N., Schmid, C. H., & Olkin, I. (2006). The case of the misleading funnel plot. *British Medical Journal, 333,* 597–600. http://dx.doi.org/10.1136/bmj.333.7568.597

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, Calif: Sage.

*Liu, A., & Want, S. (2015, June 1). *Literary fiction did not improve affective ToM*. Retrieved February 1, 2017, from http://www.PsychFileDrawer.org/replication.php?attempt=MjI1

*Mar, R. A. (2007). *Stories and the simulation of social experience neuropsychological evidence and social ramifications*. Ottawa, Ontario, Canada: University of Toronto.

Mar, R. A. (2011, July). The neural bases of social cognition and story comprehension. *Annual Review of Psychology, 62,* 103–134. http://dx.doi.org/10.1146/annurev-psych-120709-145406

Mar, R. A. (2015). *Theory and research on the relation between narrative fiction and social abilities*. Presented at the Literature and Empathy—A transdisciplinary symposium endorsed by IGEL, funded by the Thyssen foundation, Göttingen, Germany.

Mar, R. A., & Oatley, K. (2008). The Function of Fiction is the Abstraction and Simulation of Social Experience. *Perspectives on Psychological Science, 3,* 173–192. http://dx.doi.org/10.1111/j.1745-6924.2008.00073.x

Merton, R. K. (1968). The Matthew effect in science. *Science, 159,* 56–63. http://dx.doi.org/10.1126/science.159.3810.56

Mumper, M. L., & Gerrig, R. J. (2017). Leisure reading and social cognition: A meta-analysis. *Psychology of Aesthetics, Creativity, and the Arts, 11,* 109–120. http://dx.doi.org/10.1037/aca0000089

Nowicki, S., Jr., & Duke, M. P. (1994). Individual differences in the nonverbal communication of affect: The diagnostic analysis of nonverbal accuracy scale. *Journal of Nonverbal Behavior, 18,* 9–35. http://dx.doi.org/10.1007/BF02169077

Oatley, K. (2016). Fiction: Simulation of Social Worlds. *Trends in Cognitive Sciences, 20,* 618–628. http://dx.doi.org/10.1016/j.tics.2016.06.002

*Panero, M. E., Weisberg, D. S., Black, J., Goldstein, T. R., Barnes, J. L., Brownell, H., & Winner, E. (2016). Does reading a single passage of literary fiction really improve theory of mind? An attempt at replication. *Journal of Personality and Social Psychology, 111*(5), e46–e54. http://dx.doi.org/10.1037/pspa0000064

Peskin, J., & Astington, J. W. (2004). The effects of adding metacognitive language to story texts. *Cognitive Development, 19,* 253–273. http://dx.doi.org/10.1016/j.cogdev.2004.01.003

Pinkham, A. E., Penn, D. L., Green, M. F., & Harvey, P. D. (2015). Social cognition psychometric evaluation: Results of the initial psychometric study. *Schizophrenia Bulletin, 42,* 494–504.

Pinker, S. (2011). *The Better Angels of Our Nature: Why Violence Has Declined*. New York, NY: Viking.

*Pino, M. C., & Mazza, M. (2016). The use of "literary fiction" to promote mentalizing ability. *PLoS ONE, 11*(8), e0160254. http://dx.doi.org/10.1371/journal.pone.0160254

Reis, H. T., Sheldon, K. M., Gable, S. L., Roscoe, J., & Ryan, R. M. (2000). Daily well-being: The role of autonomy, competence, and relatedness. *Personality and Social Psychology Bulletin, 26,* 419–435. http://dx.doi.org/10.1177/0146167200266002

Rowe, A. D., Bullock, P. R., Polkey, C. E., & Morris, R. G. (2001). "Theory of mind" impairments and their relationship to executive functioning following frontal lobe excisions. *Brain: A Journal of Neurology, 124,* 600–616. http://dx.doi.org/10.1093/brain/124.3.600

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55,* 68–78. http://dx.doi.org/10.1037/0003-066X.55.1.68

*Samur, D., Tops, M., & Koole, S. L. (2017). Does a single session of reading literary fiction prime enhanced mentalizing performance? Four replication experiments of Kidd and Castano (2013). *Cognition and Emotion*. Advance online publication. http://dx.doi.org/10.1080/02699931.2017.1279591

Scammacca, N., Roberts, G., & Stuebing, K. K. (2014). Meta-analysis with complex research designs: Dealing with dependence from multiple measures and multiple group comparisons. *Review of Educational Research, 84,* 328–364. http://dx.doi.org/10.3102/0034654313500826

Shamay-Tsoory, S. G., & Aharon-Peretz, J. (2007). Dissociable prefrontal networks for cognitive and affective theory of mind: A lesion study. *Neuropsychologia, 45,* 3054–3067. http://dx.doi.org/10.1016/j.neuropsychologia.2007.05.021

Slaughter, V., Imuta, K., Peterson, C. C., & Henry, J. D. (2015). Meta-analysis of theory of mind and peer popularity in the preschool and early school years. *Child Development, 86,* 1159–1174. http://dx.doi.org/10.1111/cdev.12372

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21,* 360–407. http://dx.doi.org/10.1598/RRQ.21.4.1

Tamir, D. I., Bricker, A. B., Dodell-Feder, D., & Mitchell, J. P. (2016). Reading fiction and reading minds: The role of simulation in the default network. *Social Cognitive and Affective Neuroscience, 11,* 215–224. http://dx.doi.org/10.1093/scan/nsv114

Tanner-Smith, E. E., & Tipton, E. (2014). Robust variance estimation with dependent effect sizes: Practical considerations including a software tutorial in Stata and spss. *Research Synthesis Methods, 5,* 13–30. http://dx.doi.org/10.1002/jrsm.1091

Tanner-Smith, E. E., Tipton, E., & Polanin, J. R. (2016). Handling complex meta-analytic data structures using robust variance estimates: A tutorial in R. *Journal of Developmental and Life-Course Criminology, 2,* 85–112. http://dx.doi.org/10.1007/s40865-016-0026-5

Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., & Sánchez-Meca, J. (2015). Meta-analysis of multiple outcomes: A multilevel approach. *Behavior Research Methods, 47,* 1274–1294. http://dx.doi.org/10.3758/s13428-014-0527-2

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*. Advance online publication. http://dx.doi.org/10.18637/jss.v036.i03

Walberg, H. J., & Tsai, S. L. (1983). Matthew effects in education. *American Educational Research Journal, 20,* 359–373.

Yang, Y. C., Boen, C., Gerken, K., Li, T., Schorpp, K., & Harris, K. M. (2016). Social relationships and physiological determinants of longevity across the human life span. *Proceedings of the National Academy of Sciences of the United States of America, 113,* 578–583. http://dx.doi.org/10.1073/pnas.1511085112

Zaki, J., & Ochsner, K. N. (2012). The neuroscience of empathy: Progress, pitfalls and promise. *Nature Neuroscience, 15,* 675–680. http://dx.doi.org/10.1038/nn.3085

Received July 27, 2017
Revision received November 1, 2017
Accepted November 13, 2017 ■