

Psychology of Aesthetics, Creativity, and the Arts

The Impact of Fiction Reading on Social Outcomes: A 4-Week Randomized Controlled Study

David Dodell-Feder, Gloria Liu, Paige Amormino, Elizabeth D. Handley, and Diana I. Tamir

Online First Publication, October 23, 2023. <https://dx.doi.org/10.1037/aca0000617>

CITATION

Dodell-Feder, D., Liu, G., Amormino, P., Handley, E. D., & Tamir, D. I. (2023, October 23). The Impact of Fiction Reading on Social Outcomes: A 4-Week Randomized Controlled Study. *Psychology of Aesthetics, Creativity, and the Arts* Advance online publication. <https://dx.doi.org/10.1037/aca0000617>

The Impact of Fiction Reading on Social Outcomes: A 4-Week Randomized Controlled Study

David Dodell-Feder^{1, 2}, Gloria Liu¹, Paige Amormino³, Elizabeth D. Handley^{1, 4} and Diana I. Tamir⁵

¹Department of Psychology, University of Rochester

²Department of Neuroscience, University of Rochester Medical Center

³Department of Psychology, Georgetown University

⁴Mt. Hope Family Center, University of Rochester

⁵Department of Psychology, Princeton University

Does reading fiction improve our ability to understand one another? Correlational data suggest that lifetime fiction exposure is positively associated with social outcomes. Experimental data suggest that fiction reading may slightly improve social ability, although this conclusion is tenuous. Here, we test fiction's impact on social outcomes by conducting a study in which adult participants ($N = 210$) were randomly assigned to engage in no reading for pleasure, or to read fiction or nonfiction for 45 min/day, 5 days/week, for 4 weeks. At the end of the study, participants were assessed on three classes of social outcomes: theory of mind, empathy, and social functioning. Using structural equation modeling, we tested the impact of fiction reading on latent variables representing the aforementioned social outcomes. Fiction readers did not outperform nonfiction readers or participants who abstained from pleasure reading on any social outcome. Nonfiction readers outperformed those who abstained from pleasure reading on the empathy latent variable. We did not observe associations between lifetime fiction exposure and social outcomes. Taken with the study's limitations, which include a modest sample size, measurement issues, and the possibility that non-study media was consumed/produced during the reading period, these data are consistent with the following possibilities regarding fiction's positive social impact: Such findings may reflect a priming effect, may occur only after prolonged exposure to fiction, and/or may occur for readers who exhibit a particular kind of engagement with the reading. Together, this study provides no new evidence of a beneficial effect of fiction reading on social outcomes.

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

Supplemental materials: <https://doi.org/10.1037/aca0000617.supp>

Can stories help people to understand one another? Scholars across time and disciplines have long thought the answer to be “yes” (Black et al., 2021). After all, stories and their telling in the context of narrative fiction typically focus on people, their interactions, and their emotions (Hogan, 2003; Mar & Oatley, 2008; Schwering et al., 2021), inducing embodied simulation (Mar, 2004; Mar & Oatley, 2008; Oatley, 1999, 2016; Willems & Jacobs, 2016; Zwaan, 2016). Scholars have gone so far as to describe fiction as “the mind’s flight simulator” (Oatley, 2016, p. 619) in which fiction readers recruit the

very same processes used to understand people in real life (Gerrig, 1993; Mar, 2018b; Mar & Oatley, 2008; Oatley, 1999, 2016), a notion supported by neuroimaging research (Hartung & Willems, 2020; Mar, 2004, 2011; Mason & Just, 2009; Nijhof et al., 2015; Tamir et al., 2016; Willems & Hartung, 2017). As described in the Social Processes and Content Entrained by Narrative framework (Mar, 2018b) and elsewhere (Oatley, 2016), repeated exposure to fiction may confer the hypothesized beneficial social effects through (a) the process of repeated social simulation as a form of social practice

David Dodell-Feder  <https://orcid.org/0000-0003-0678-1728>

This project was generously supported by an award from the Research: Art Works program at the National Endowment for the Arts (Grant 1809928-38-18) and startup funds provided by the University of Rochester. The opinions expressed in this paper are those of the authors and do not represent the views of the Office of Research and Analysis or the National Endowment for the Arts. The National Endowment for the Arts does not guarantee the accuracy or completeness of the information included in this paper and is not responsible for any consequence of its use.

The authors would like to thank Christine I. Hooker for feedback regarding the study design, members of the Rochester Social Cognition and Psychopathology Laboratory for feedback on the manuscript, and the

following people for assistance with data collection, data entry, and administrative management: Aly Chenaille, Anita Kwashie, Anne Zhou, Arisha Ansari, Ayde Amir, Chenyuan Qu, Cheryl Best, Christie Ulloa, Dalya Hahn, Elizabeth Schupbach, Emily Dudek, Emma Saubermann, Gianna C. Geil, Grace Glover, Isabel Adeniji, Jessica Woodyatt, Kailin Zhuang, Katherine Wang, Lia Rodriguez, Mars Scharf, Matthew Yung, Nicole Calcagno, Nya Robinson, Rachel Bordonaro, Runako Campbell, Sanoelani Lanais, Sara Signorelli, Shangzan Liu, Siddhant Iyer, Stephanie Reda, Tiffany Chen, and Tony Phan. The authors would also like to thank all of the participants.

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

and (b) the social content of the reading that provides readers with concrete social information about the world, which readers can then apply to their advantage in real-world social interaction.

If scholars are correct, and fiction reading does improve social abilities, the consequences would be profound. For example, in the United States, empathy is on the decline (Konrath et al., 2011), contributing to a widespread “empathy deficit” (Hall & Leary, 2020), which may in turn be reducing subjective well-being and civility (Chopik et al., 2017). Meanwhile, political polarization is rising in the United States (Boxell et al., 2020; Finkel et al., 2020), with a concomitant rise in outgroup hate that may actively promote dehumanization and hamper motivation to understand one another. Social cognitive impairment continues to be a transdiagnostic problem in psychiatric and neurological disorders (Cotter et al., 2018), contributing to the daily functioning difficulties that make these disorders emotionally and economically costly for the individual and society. Could fiction reading be part of the solution to these problems?

Researchers have conducted both correlational and experimental studies to answer this question. Correlational studies have shown that increased lifetime fiction reading is associated with improved performance on self-report and performance-based measures of (a) theory of mind (ToM; Black & Barnes, 2015a; Black & Barnes, 2015b; Fong et al., 2013; Kidd & Castano, 2013; Mar et al., 2006, 2009; Panero et al., 2016; Tamir et al., 2016)—that is, the ability to attribute and reason about the mental states of others—(b) empathy (Mar et al., 2006, 2009)—that is, considering and vicariously sharing another person’s internal states—and (c) social functioning (Mar et al., 2009)—that is, loneliness, social network size, etc.—even after controlling for potentially confounding variables (e.g., Mar et al., 2009). A meta-analysis of this literature (Mumper & Gerrig, 2017) demonstrated a medium-sized association between lifetime fiction reading and ToM, and small- to medium-sized associations between lifetime fiction reading and empathy, supporting the idea that fiction reading is associated with social ability. In further support of a fiction reading and social outcome connection, longitudinal data also suggest an association over time between reading comprehension and later ToM ability in children (Lecce et al., 2021; although see van der Kleij et al., 2022), reading for pleasure and later emotional adjustment and prosocial behavior in children (Mak & Fancourt, 2020), and fiction reading and later helping behavior in adults (Kou et al., 2020). The benefit of approaching fiction’s putative positive impact through correlational work is that these studies may speak to how the fiction–social ability association operates in real life, with readers having agency over when, how much, and what they read, in a natural context, unlike a laboratory setting (Black et al., 2021). The downside is that these data cannot determine whether the association is causal in nature. Does fiction reading improve social outcomes, or do people with stronger social skills choose to read more fiction?

Experimental work, in contrast, can speak to the possible causal effects of fiction reading. Indeed, experiments have shown that after reading fiction, people show greater ToM performance (Kidd & Castano, 2013, 2019; Kidd et al., 2016; Pino & Mazza, 2016), empathy (Bal et al., 2013; Djikic et al., 2013; Johnson et al., 2013), and prosocial behavior (Johnson et al., 2013; Koopman, 2015) compared to nonfiction or no reading. This occurs even when the exposure to fiction is short (e.g., a single short story). There are, however, notable exceptions to these findings. Several careful and well-powered replication attempts have found no impact of fiction reading on social outcomes (Lenhart & Richter, 2022; Panero et al., 2016; Samur et al.,

2018). A meta-analysis across all of these disparate findings in the experimental literature demonstrated a positive, albeit small, benefit to fiction reading (Dodell-Feder & Tamir, 2018). Subsequent *p*-curve analysis has shown that this body of work has evidential value, meaning that it is unlikely to have been influenced by questionable research practices (Quinlan et al., 2023). That said, this conclusion was also not robust, since the evidential value becomes inconclusive when dropping single studies from the analysis. In sum, though there exists evidence that fiction reading causally improves social outcomes to a small degree, this conclusion is tenuous at best.

Furthermore, there are several limitations to the existing literature on the possible causal effects of fiction reading on social outcomes (Mar, 2018a, 2018b). First, fiction’s effects have been overwhelmingly assessed with a small number of social outcome measures (Dodell-Feder & Tamir, 2018; Mumper & Gerrig, 2017). Most often, social cognition is measured with the Reading the Mind in the Eyes Task (RMET; Baron-Cohen et al., 2001), a performance-based measure of the ability to decode mental states from solely the eyes of White actors, and the Interpersonal Reactivity Index (IRI; Davis, 1983), a self-report measure of multiple dimensions of empathy. Though both measures are widely used in the social cognitive literature, and exhibit adequate psychometric properties (Davis, 1983; Olderbak et al., 2015; cf., Black, 2019 and Dodell-Feder et al., 2020), these two measures do not adequately assess the full suite of processes involved in ToM and empathy, respectively. Furthermore, besides a few notable exceptions (Johnson, 2012; Johnson et al., 2013; Koopman, 2015; Kou et al., 2020; Mar et al., 2009), there have been few attempts to evaluate whether fiction’s positive impact extend to the downstream consequences of social cognition, into the domain of real-world interpersonal effectiveness and social functioning. As such, it is possible that fiction’s positive impact is largely measure-specific, and does not generalize to a broader category of social outcomes.

Second, the overwhelming amount of experimental work involves short exposures to fiction (cf., Pino & Mazza, 2016). In contrast, real-world reading habits involve frequent, prolonged, and intense engagement with reading (Perrin, 2016; The Rise of E-Reading, 2012). It is possible that a short fiction exposure may indeed produce a real, lasting social change (i.e., the meta-analytic $g = 0.15\text{--}0.16$ found in Dodell-Feder & Tamir, 2018). However, these effects may underestimate the impact of real-world fiction reading. When the effects of reading have the chance to accumulate over more extended time periods, after more intense exposures, and with voluntary control over the reading material, this may lead to larger and more practically significant effects than the ones observed thus far. Alternatively, short exposure to fiction in a laboratory may produce a positive but short-lasting impact on social outcomes that cannot be accounted for by real change to social cognition, but instead by a priming effect. If so, we would not expect longer, more intense reading experiences to offer more robust or more significant evidence for a causal effect of reading on social outcomes.

We conducted a longitudinal randomized controlled study to evaluate fiction’s putative causal impact on a broad set of social outcomes to arbitrate between these possibilities. Participants were randomized to engage in no reading for pleasure, or to read fiction or nonfiction for 45 min/day, 5 days/week, for 4 weeks. Fiction material included classic literary fiction used in other experiments; nonfiction material was nonsocial (e.g., texts on natural science topics). Participants were free

to read any of the large number of texts we provided. We selected outcome measures to sample three related but distinct major facets of social processes supported by multiple streams of behavioral, neuroimaging, and clinical research (Arioli et al., 2021; M. F. Green et al., 2015; Ochsner, 2008; Schurz et al., 2021; Zaki & Ochsner, 2012): ToM, empathy, and social functioning. Separating and modeling social outcomes in this manner is also consistent with how fiction's impact on social outcomes has been studied in prior work (Dodell-Feder & Tamir, 2018; Mumper & Gerrig, 2017). We selected specific measures used in prior studies of fiction's impact on social outcomes (e.g., RMET, IRI, Moral Judgment Task [MJT], Social Network Index [SNI], University of California, Los Angeles Loneliness Survey; see below) in addition to measures not previously used in such experiments in an effort to broadly sample each construct. Prior work has shown that a reader's experience with fiction influences whether fiction impacts social outcomes (Bal et al., 2013; Johnson, 2012; Johnson et al., 2013; Schwerin & Lenhart, 2022). To test this possibility, we included measures of narrative transportation to assess the extent to which one becomes mentally and emotionally absorbed into the narrative, and intrinsic motivation to assess the extent to which one experienced enjoyment, agency, and value in reading during the study. Finally, we included a measure of lifetime exposure to fiction—the widely used Author Recognition Test (ART) (Mar et al., 2006)—to evaluate previously documented associations between lifetime exposure to fiction and social outcomes (Mumper & Gerrig, 2017), and test a possible interactive effect of lifetime fiction reading with an intensive fiction reading experience on social outcomes (Kidd & Castano, 2019). We used structural equation modeling (SEM) to test the following hypotheses: (a) The fiction reading group will outperform the nonfiction and no pleasure reading group (hereafter, “no reading”) on the social outcomes; (b) lifetime exposure to fiction will be positively associated with social outcomes; (c) lifetime exposure to fiction will influence whether the intensive fiction reading experience impacts the social outcomes; and (d) greater narrative transportation and intrinsic motivation will be associated with better performance on the social outcomes in the fiction group.

Method

Open Science Practices

The study was preregistered on the Open Science Framework at the following link: <https://doi.org/10.17605/OSF.IO/7MFWJ>. The analysis plan and target sample size were preregistered shortly after starting data collection but prior to any analyses. In our preregistration, we included a broad set of measures assessing both social and nonsocial processes that may be influenced by fiction reading. The focus of the current paper was on social outcomes for which there exists the strongest evidence that fiction may positively impact (Dodell-Feder & Tamir, 2018; Mumper & Gerrig, 2017). Thus, we omitted all nonsocial measures collected as part of the overall study from analysis. However, those data are freely available from the authors upon request (see the online supplemental materials for a list of other measures not analyzed here). The *social stroop* task included in the preregistration was not analyzed due to an error in the stimulus presentation code that made the data unusable. Likewise, of the 210 participants with complete datasets, *pencil drop prosocial behavior task* data were missing for a substantial number (36%) of participants and so we did not include these data in the analysis. Data and analysis

code are available on the Open Science Framework at https://osf.io/6pczf/?view_only=dacf6af4faf44b3bb45ee4f41de922e4.

Sample Size Determination and Power Analysis

A target sample size ($N = 300$) was determined based on the planned analytic strategy described in our preregistration. However, after discussion among the authors since the preregistration was completed in 2017, and prior to conducting any analyses, we determined that a different analytical approach was more appropriate. Namely, we used SEM to take advantage of our study's multivariate, confirmatory nature, focused on latent constructs (i.e., ToM, empathy, social functioning) comprised of multiple related measures. SEM also offered better control of Type I error rate than our original analytic plan. We provide the results from analyses that more closely resemble our preregistered analysis plan in the online supplemental materials. Importantly, the results from these analyses align with those provided by SEM.

We used the *semPower* package (Jobst et al., 2023) and *WebPower* (Zhang & Mai, 2023) to perform sensitivity analyses. Given N and df for each model described below, the probability of rejecting a misspecified model with a corresponding standard model fit criterion of root-mean-square error of approximation (RMSEA) ≥ 0.06 (Hu & Bentler, 1999) across the models was $\geq 96\%$ ($\alpha = .05$). We note though that our N is modest according to most sample size rules-of-thumb for SEM (Kyriazos, 2018), and, more generally, that variable associations stabilize at N s approaching 250 (Schönbrodt & Perugini, 2013). Thus, findings from our models should be interpreted with caution. For the regression models, given N and $\alpha = .05$, we had 80% power to detect effects of $f^2 = 0.06$, or small- to medium-sized effects.

Participants

A total of 287 individuals were enrolled in the study. Of those individuals, 57 participants (19.9%) withdrew from the study before their final study visit due to not being able to keep up with the reading schedule, failure to respond to study personnel, and/or lockdown mandates during the COVID-19 pandemic that prevented participants from returning to the lab for the postreading assessment, leaving a final sample of 230 participants who completed the study. There were no differences in demographic characteristics between subjects who withdrew from and completed the study. Due to data loss that occurred while transitioning the project to another institution and technical errors, of those 230 participants, we had complete outcome data on the social measures from 210 participants, and outcome data plus the covariates used in the models described below from 167 to 207 participants. We note N in all models described in the Results section.

Participants were recruited from local universities and the community. Inclusion criteria included being between the ages of 18–25 (i.e., approximately college-aged), fluent in English, and having access to an internet-connected device in order to access the reading material and complete web-based reading logs. Exclusion criteria included being diagnosed with a psychiatric, learning, or neurological disorder. Demographic information was missing for 53 individuals either because the individual chose not to report that information or due to data loss.

Of those participants whose demographic information was available, participants were on average in their early 20s, female, White,

Table 1
Demographic Information and Reading Data

| Variable | Fiction | Nonfiction | No reading control | Group comparison |
|---|-------------------|-------------------|--------------------|---|
| <i>n</i> ^a | 71 | 63 | 76 | |
| Age, <i>M</i> (<i>SD</i>) | 21.2 (1.9) | 21.2 (2.1) | 21.2 (2.1) | $F(2, 172) = 0.00, p = .997, \eta^2 < .001$ |
| Sex, <i>n</i> (%) | | | | $\chi^2(2, N = 174) = 2.28, p = .320, V = 0.11$ |
| Female | 45 (78) | 34 (65) | 48 (75) | |
| Male | 13 (22) | 18 (35) | 16 (25) | |
| Race, <i>n</i> (%) | | | | $\chi^2(6, N = 176) = 6.45, p = .375, V = 0.14$ |
| Asian | 24 (41) | 15 (28) | 16 (25) | |
| Black | 4 (7) | 3 (6) | 8 (12) | |
| White | 26 (44) | 32 (60) | 36 (56) | |
| Mixed/other | 5 (8) | 3 (6) | 4 (6) | |
| Ethnicity, <i>n</i> (%) | | | | $\chi^2(2, N = 177) = 1.32, p = .517, V = 0.09$ |
| Hispanic/Latinx | 4 (7) | 3 (6) | 7 (11) | |
| Non-Hispanic/Latinx | 55 (93) | 51 (94) | 57 (89) | |
| Education (years), <i>M</i> (<i>SD</i>) | 14.1 (2.2) | 14.5 (1.9) | 13.6 (2.9) | $F(2, 113) = 1.29, p = .280, \eta^2 = .022$ |
| Wide Range Achievement Test composite, <i>M</i> (<i>SD</i>) | 115.9 (11.4) | 116.8 (9.6) | 114.9 (9.5) | $F(2, 206) = 0.59, p = .558, \eta^2 = .006$ |
| Days read, <i>M</i> (<i>SD</i>) | 19.3 (4.8) | 20.2 (5.1) | | $IRR = 1.04, 95\% \text{ CI } [0.96, 1.13], p = .328, R^2 = .007$ |
| Minutes/day read, estimated marginal <i>M</i> [95% CI] | 47.7 [45.5, 49.9] | 46.3 [44.0, 48.6] | | $b = -1.39, t = 0.87, p = .388, R^2 = .002$ |
| Reading enjoyment, estimated marginal <i>M</i> [95% CI] | 3.5 [3.4, 3.7] | 3.4 [3.3, 3.6] | | $b = -0.11, t = 0.90, p = .370, R^2 = .003$ |
| Percent accuracy on attention/comprehension questions, estimated marginal <i>M</i> [95% CI] | 85.8 [84.3, 87.3] | 89.7 [88.1, 91.3] | | $b = 3.92, t = 3.53, p < .001, R^2 = .010$ |
| Days of nonstudy-related leisure reading, <i>n</i> (%) | 65 (5) | 25 (2) | | $IRR = 0.54, 95\% \text{ CI } [0.15, 1.99], p = .353, R^2 = .008$ |
| Narrative Transportation Scale, <i>M</i> (<i>SD</i>) ^b | 4.3 (0.7) | 3.7 (0.8) | | $t(121) = 4.43, p < .001, d = 0.78 [0.43, 1.14]$ |
| Intrinsic Motivation Inventory, <i>M</i> (<i>SD</i>) ^c | 5.2 (0.8) | 5.1 (0.9) | | $t(120) = 0.73, p = .467, d = 0.13 [-0.21, 0.47]$ |

Note. Data are from participants with complete outcome variable data ($N = 210$). Totals may be less than 210 due to missing data. Estimated marginal means are derived from linear mixed-effects models that included a random intercept for participant. Fiction is the reference group in regression models. IRR = Incidence Rate Ratio; CI = confidence interval.

^a *n* refers to number of cases with complete outcome variable per group. ^b Fiction $n = 69$, nonfiction $n = 61$. ^c Fiction $n = 70$, nonfiction $n = 62$.

non-Hispanic/Latinx, and had 14 years of education (Table 1). There were no differences among these characteristics between groups. Reading composite scores on the Wide Range Achievement Test (WRAT; Wilkinson & Robertson, 2006) indicated above-average reading ability. Similarly, there were no differences between groups on the WRAT composite score.

Participants were compensated monetarily for their participation at \$10 per hour for the behavioral sessions and \$5 per day for reading, and were given additional monetary bonuses of \$5 per week of reading on schedule. The study was approved by the Princeton University Institutional Review Board and the University of Rochester Research Subjects Review Board.

Procedure

The study was completed in three phases. In Phase 1, participants completed a prereading assessment in lab during which they provided demographic information, completed the WRAT, and were assessed on the outcome measures described below. The experimenter who performed the assessments was blind to the participant's group assignment at the pre- and postreading assessment. At the end of the first session, an independent experimenter assigned participants to either fiction reading, nonfiction reading, or no reading. The independent experimenter oriented the participant to the reading protocol and followed their progress over the next four weeks. Participants only knew whether they were assigned to a reading or no reading group, but did not know about the existence of a second reading group (i.e., fiction or nonfiction), nor the study hypotheses.

In Phase 2, participants in the fiction and nonfiction reading group were given electronic material to read according to a schedule of 5 days/week, 45 min/day for the 4 weeks following their initial study visit. Participants were free to read any days of the week and at any time during the day. Fiction material consisted of short stories and novellas generally considered classic literary fiction (e.g., *The Chameleon* by Anton Chekhov; see the online supplemental materials). We chose to include as much material as we could access that has been used in other work evaluating fiction's impact on social cognition. Nonfiction material consisted of short works primarily on natural science topics (e.g., astronomy) that were nonsocial. Biography, memoir, and other material focused on people or social relationships were not included. This approach, while conflating fictionality with social content, is typical of work in the field. However, we note that it prevents us from making inferences regarding the causal impact of fictionality versus social content, which has been shown to make a difference on social outcomes in other mediums (i.e., writing, as in Black & Barnes, 2021). Fiction and nonfiction readers were asked not to read for pleasure outside of the material provided to them as part of the study. No reading participants were asked to not engage in any pleasure reading during the study period.

Each day a participant read, they completed a Daily Reading Questionnaire where they reported the material they read, how long they read, and whether they read outside material not part of the study for pleasure. In addition, for each reading selection completed, participants were asked "How much did you enjoy reading this?" which they responded to using a continuous 0 (*not at all*)

to 5 (*extremely*) star rating scale, and a set of questions assessing their attention to the specific reading selection. Daily Reading Questionnaire responses were monitored by study staff to ensure adherence to the protocol. If participants missed 2 days of reading, they were prompted by study staff to ensure they stayed on schedule. Participants in the fiction reading condition were sent an additional questionnaire once per week (four in total) assessing narrative transportation (see below). After the 4-week period, participants were brought back into the lab for the postreading assessment.

In Phase 3, participants completed the postreading assessment in which they completed the same set of tasks prereading. In addition to these measures, all participants completed a behavioral measure of prosocial behavior (see the online supplemental materials), and participants in the fiction and nonfiction groups completed a measure of intrinsic motivation (see below). After completing all study measures and receiving payment, participants completed a Reading Fidelity Questionnaire anonymously, where they reported whether they deviated from the reading schedule in any way. We used this questionnaire to assess general fidelity to the reading protocol. Due to experimenter error, these data were collected for 128 out of 149 participants assigned to read either fiction or nonfiction.

Measures

Outcome measures were selected to sample ToM, empathy, and social functioning. Separating and modeling social outcomes in this manner is consistent with how fiction's impact on social outcomes has been studied in prior work (Dodell-Feder & Tamir, 2018; Mumper & Gerrig, 2017). We made an effort to include measures used in prior work testing fiction's impact on social cognition as well as measures that assess previously untested social outcomes. Below we describe only those measures included in the final models; all other measures are described in the online supplemental materials and our preregistration.

ToM Outcome Measures

RMET. The RMET (Baron-Cohen et al., 2001) assesses an individual's ability to decode mental and affective states depicted by solely the eye regions of White actors/actresses by selecting which of four adjectives (e.g., "joking," "insisting," "amused," "relaxed") best describes the mental/affective state expressed by the eyes. Participants complete one practice trial and 36 experimental trials. Given the complexity of the adjectives used as response options (Dodell-Feder et al., 2020), participants are provided with a dictionary that defines each response option. Scores reflect the number of correct trials. The RMET is the most commonly used measure to assess ToM in the fiction reading literature (Dodell-Feder & Tamir, 2018; Mumper & Gerrig, 2017). In our sample, omega was .67, 95% confidence interval (CI) [0.43, 0.81], which is similar to other fiction reading studies using the RMET (e.g., $\alpha = .68$ in van Kuijk et al., 2018) and a recent meta-analysis of the psychometric properties of the RMET ($\alpha = .73$ in Kittel et al., 2022).

Emotion Identification Task. In this measure of emotion identification (Germine & Hooker, 2011), participants view pictures of one White male and two White female faces. Each face is morphed between a neutral face and one of four emotional

expressions: happy, disgusted, angry, and fearful. Participants complete 60 trials in which they saw a face and identified which of the four emotions it was expressing. Scores reflect the percentage correctly identified.

MJT. The MJT (Young et al., 2012) assesses the degree to which a participant considers a character's intention when judging the moral permissibility of the character's actions. Participants read 24 short vignettes in a two intention (negative, neutral) by two outcome (negative, neutral) design. We were interested in the two cases when intention and outcome conflict, namely, accidental harm scenarios (MJT-Acc) in which a character causes harm despite having a neutral intention, and attempted harm scenarios (MJT-Att) in which a character does not cause harm despite having a negative intention. After each vignette, participants rate the moral permissibility of the character's action using a 1 (*morally forbidden*) to 7 (*morally permissible*) scale. Brain regions selective for ToM show preferential activation for these scenarios in which the intention differs from the outcome (Young et al., 2007), and disruption to these regions reduces the role of mental state information when judging moral permissibility (Young et al., 2010). In addition, individuals with ToM impairments show altered ratings for these scenarios (Moran et al., 2011; Young et al., 2012). Thus, to the extent that fiction reading improves ToM, we would expect that, compared to the control groups, ratings of moral permissibility would be higher in the accidental harm scenarios and lower in the attempted harm scenarios, reflecting greater attention to the character's intention versus the outcome of an action. Scores on the MJ-Att were reversed so that as in the MJ-Acc condition, higher scores indicated greater consideration of intention.

Empathy Outcome Measures

IRI. The IRI (Davis, 1983) is a self-reported multidimensional scale of empathy. In the current study, participants completed all four IRI subscales with three of the four being included in the final model: perspective-taking (IRI-PT), which assesses the extent to which an individual considers others' point-of-view (e.g., "I sometimes find it difficult to see things from the other guy's point of view"; reversed-scored), empathic concern (IRI-EC), which assesses emotional concern for others (e.g., "I often have tender, concerned feelings for people less fortunate than me"), and fantasy (IRI-FS), which assesses the tendency to identify with and become absorbed into fictional characters depicted in books, movies, or other media (e.g., "I really get involved with the feelings of the characters in a novel"). Each dimension is assessed with seven items rated using a 0 (*does not describe me well*) to 4 (*describes me very well*) scale, with higher scores indicating greater empathy. In our sample, omega was .85, 95% CI [0.81, 0.88] for IRI-PT, .83, [0.78, 0.87] for IRI-EC, and .85, [0.81, 0.89] for IRI-FS.

Balanced Emotional Empathy Scale (BEES). The BEES (Mehrabian & Epstein, 1972) is a self-report measure of the vicarious experience of another's emotional experience. We used the abbreviated seven-item version of the questionnaire. Participants responded to each question using a -4 (*very strong disagreement*) to $+4$ (*very strong agreement*) scale, with higher scores indicating greater empathy. In our sample, omega was .65, 95% CI [0.58, 0.72], which is similar to the value reported in other studies (Mehrabian, 2000).

Social Functioning Outcome Measures

SNI. The SNI (Cohen, 1997) is a self-report measure that assesses the size and nature of one's social network by evaluating one's participation in 12 social roles (e.g., spouse, parent, close friend). We calculated two scores: (a) the number of people in one's social network in which one has regular contact (SNI-P) and (b) the number of social roles in which one has regular contact with others, referred to as "network diversity" (SNI-ND).

UCLA Loneliness Scale (ULS). The ULS (Russell, 1996) is a 20-item self-report measure that assesses loneliness (e.g., "How often do you feel that you lack companionship?"). Questions are answered using a 1 (*never*) to 4 (*always*) scale, with higher scores indicating greater loneliness. Scores were reversed to be consistent with the other social functioning measures, such that higher scores indicated less loneliness. In our sample, omega was .94, 95% CI [0.93, 0.95].

Medical Outcomes Study—Social Support Survey (MOS-SSS). The MOS-SSS (Sherbourne & Stewart, 1991) is a 19-item self-report questionnaire that assesses how often various kinds of support (i.e., emotional, tangible, affectionate, and positive social interaction support) are available to participants (e.g., "Someone to help you if you were confined to bed"). Questions are answered on a 1 (*none of the time*) to 5 (*all of the time*) scale, with higher scores representing more social support. In our sample, omega was .95, 95% CI [0.94, 0.96].

Other Measures

Narrative Transportation Scale (NTS). The NTS (M. C. Green & Brock, 2000) is a 12-item self-report questionnaire that assesses absorption into a story (e.g., "I was mentally involved in the narrative while reading it"). Questions are answered on a 1 (*not at all*) to 7 (*very much*) scale, with total scores representing the sum of the items such that greater scores indicate greater transportation. The NTS was administered once per week to reading fiction and nonfiction participants. Given multiple entries per participant, we calculated multilevel omega. Within and between omega was .56, 95% CI [0.49, 0.64], and .87, [0.83, 0.90], respectively.

Intrinsic Motivation Inventory (IMI)

The IMI (Choi et al., 2010) is a 21-item self-report question that assesses interest/enjoyment, perceived choice, and value/usefulness of a given activity. Questions are answered with a 1 (*not at all*) to 7 (*very true*) scale, with higher scores indicating greater intrinsic motivation. Fiction and nonfiction participants completed this questionnaire only during the postreading assessment session. In our sample, omega was .93, 95% CI [0.90, 0.95].

ART

The ART (Mar et al., 2006) is a measure of exposure to different types of print. The ART taps into real-world reading intentions (Rain & Mar, 2014) and real-world reading habits (West et al., 1993). Participants are instructed to indicate which authors they recognize from a list of fiction authors, nonfiction authors, and foils (names of individuals who are not either), which are included to discourage guessing. Lifetime exposure to fiction is calculated by taking the

number of fiction authors recognized by the participant and subtracting the number of foils to account for guessing. Higher scores indicate higher lifetime exposure to fiction. Prior work has shown that higher scores on the ART are associated with higher scores on measures of ToM and empathy (Mumper & Gerrig, 2017). Due to the way we collected the data, we were unable to analyze item-level data and calculate omega, although other studies using this version of the ART have found that the fiction scale exhibits internal consistency (Cronbach's α) values $\geq .90$ (Mar et al., 2006, 2009).

Data Analysis

All data were analyzed in R statistical software (R Core Team, 2022) and R Studio (RStudio Team, 2020) using the following packages: *confintr* (Mayer, 2022), *emmeans* (Lenth, 2022), *Hmisc* (Harrell, 2022), *lavaan* (Rosseel, 2012), *lmer4* (Bates et al., 2015), *lmerTest* (Kuznetsova et al., 2017), *MASS* (Venables et al., 2002), *MBESS* (Ken, 2022), *multilevelTools* (Wiley, 2020), *performance* (Lüdtke et al., 2021), *psych* (Revelle, 2018), *rstatix* (Kassambara, 2021), and *semTools* (Jorgensen et al., 2018).

First, we conducted a confirmatory factor analysis (CFA) to identify the latent factor structure of the social outcome variables, separately for each time point. We iterated this process until we identified a measurement model with acceptable fit (see below) with all indicator variables exhibiting acceptable loadings on the latent factor (standardized loadings ≥ 0.30). Using the resulting three-factor measurement model, comprised ToM, empathy, and social functioning, next, we conducted a series of SEMs to evaluate our main study questions. We intended to control for prereading scores using the latent variables derived from the CFA. However, individual indicators demonstrated strong stability from pre- to postreading with correlations as high as $r = .81$ for some measures. Furthermore, the latent factors did not vary by group at baseline (the online supplemental materials) suggesting that the random assignment was successful. Consequently, all analyses were conducted using postreading data only.

To address our main study questions, we conducted the following structural models with the latent variables as the outcome: In the first model, group assignment was the predictor, which was dummy-coded with fiction as the reference group; in the second model, group assignment and lifetime fiction exposure (ART score) were the predictors; in the third model, group assignment, lifetime fiction exposure, and their interaction (calculated as the product of the indicators) were the predictors. We reran models with nonfiction as the reference group to test for differences in the structural paths between nonfiction and no reading. As demographic variables have been shown to impact scores on social cognitive measures (Dodell-Feder et al., 2020), and reading ability may impact the extent to which participants were able to engage with the reading material, we also reran all models controlling for demographic variables (age, sex, race/ethnicity coded as White/non-Hispanic/Latinx or non-White and/or Hispanic/Latinx) and reading ability (WRAT composite score).

Models were estimated using robust maximum likelihood estimation, which generated statistics robust to nonnormal distributions in the data, such as skewness and kurtosis. As RMET scores were negatively skewed, which is common with this measure (Black, 2019), we applied a square root transformation to the scores and reran the measurement model described above and main SEMs (with group and then group and ART as the predictor). Findings were unchanged and so we report models using the untransformed data in the

manuscript (see the R Markdown file for the analyses using transformed data). We modeled residual covariances between the SNI-P and SNI-ND as these scores were directly related (i.e., SNI-ND is determined, in part, by SNI-P). Acceptable model fit for the CFAs and SEMs were evaluated using the following criteria: nonsignificant ($p > .05$) χ^2 statistic, Comparative Fit Index (CFI) values >0.95 , RMSEA values <0.06 , and standardized root mean square residual (SRMR) values <0.08 (Hu & Bentler, 1999).

We tested our secondary hypothesis that narrative transportation and intrinsic motivation might be related to whether fiction reading impacted the latent variables. Given that these questions pertained to the reading groups specifically, and the analyses would have a substantially reduced N that would make SEM inappropriate, we generated factor scores for the latent variables and then conducted two sets of multiple linear regressions, one for narrative transportation using M NTS across the multiple reports during the reading period, and one for intrinsic motivation using the IMI score. Specifically, first, we regressed each latent variable score on the reader variable (narrative transportation, and, in separate models, intrinsic motivation). Next, we regressed each latent variable score on the reader variable, condition, and the interaction between these terms. If reader engagement determines whether fiction reading impacts the social outcome, we should observe a condition by reader variable interaction, such that increasing levels of narrative transportation and intrinsic motivation would be associated with greater social outcomes for the fiction group specifically.

Results

Reading Data

Overall, participants adhered closely to the reading protocol, reading on average for 19.7 out of the 20 prescribed days and for 47.1 min/day out of the 45 prescribed minutes. There were no differences between groups in either value (Table 1). Participants in the fiction and nonfiction groups reported reading material for pleasure that was not assigned as part of the study in only 3% of days. In the fiction and nonfiction group, responses from the Reading Fidelity Questionnaire ($N = 128$) indicated that 88% of participants reported reading logs accurately in terms of whether or not they read, 8% of participants reported reading one to two times when they did not, and 5% reported three to four times reading when they did not. Participants rarely underreported the amount of time spent reading. Eighty one percent of participants reported never reading for a longer time than what was reported, 5% reported doing this rarely, 9% reported doing this sometimes, and 5% reported doing this often. Of the 19 participants who provided information about reporting discrepancies, the majority (58%) noted reading 1–10 min less than what they reported. Comprehension of and attention to the readings was adequate, with participants answering questions with 87.6% accuracy. Accuracy was higher in the nonfiction group ($p < .001$), although the difference was small ($R^2 = .01$), and may have been a function of the fiction-reading questions being more difficult, as some questions asked about nonfactual information, requiring participants to make inferences about characters and their interactions. Participants reported moderate enjoyment of the readings, giving them, on average 3.5 out of five stars, which did not differ between groups. Thus, overall, we have good reason to believe that participants performed the reading largely as prescribed. In those instances

where participants deviated, the degree of deviation was small and did not differ across groups.

Descriptive Information and Measurement Model

Descriptive statistics and bivariate correlations between study variables are presented in Tables 2 and 3, respectively (descriptive statistics for each group are provided in the online supplemental materials). Within each type of measure—ToM, empathy, social functioning—study variables were positively correlated and ranged from small-to-large in magnitude (using benchmarks described in Gignac & Szodorai, 2016) for the empathy measures, $r_s = .08-.77$, and medium-to-large for the ToM measures, $r_s = .22-.39$, and social functioning measures, $r_s = .26-.91$.

We iteratively refined the measurement model with CFA until determining an appropriate factor structure. The resulting three-factor model—ToM, empathy, social functioning—demonstrated good fit to the data: $\chi^2(50) = 55.77$, $p = .267$, CFI = 0.99, RMSEA = 0.02 (90% CI [0.00, 0.05]), SRMR = 0.05. All factor loadings were ≥ 0.35 and statistically significant ($p < .001$; Table 3).

Performance on the individual outcome measures by group at postreading is presented in the online supplemental materials. Between-group differences were unexpected under the null hypothesis ($p < .05$) and small in magnitude ($\eta^2 = .03-.05$) for MJ-Att and BEES. The nonfiction group outperformed the fiction and no reading group on MJ-Att ($d = 0.55$, 95% CI [0.21, 0.87], $d = 0.48$, [0.14, 0.81], respectively, Tukey Honestly Significant Difference (HSD)

Table 2

Descriptive Statistics for Measured Variables

| Variable | <i>M</i> | <i>SD</i> | Min–max | Skew | Kurtosis |
|--------------------------|----------|-----------|-----------|-------|----------|
| RMET | 29.4 | 3.7 | 11–35 | –1.69 | 5.15 |
| EIT | 0.69 | 0.09 | 0.30–0.93 | –0.66 | 1.30 |
| MJ-Acc | 4.8 | 1.2 | 1–7 | –0.49 | –0.29 |
| MJ-Att ^a | 5.6 | 1.0 | 3–7 | –0.80 | 0.41 |
| IRI-PT | 19.9 | 5.3 | 3–28 | –0.47 | –0.21 |
| IRI-EC | 19.8 | 5.4 | 1–28 | –0.83 | 0.70 |
| IRI-FS | 19.8 | 6.0 | 0–28 | –0.63 | –0.13 |
| BEES | 11.3 | 8.4 | –12–28 | –0.46 | –0.17 |
| SNI-P | 19.6 | 9.0 | 3–56 | 0.49 | 0.58 |
| SNI-ND | 2.3 | 1.4 | 0–7 | 0.18 | –0.50 |
| ULS ^b | 38.6 | 11.3 | 5–65 | –0.49 | 0.13 |
| MOS-SSS | 4.0 | 0.8 | 1.4–5.0 | –0.68 | –0.15 |
| NTS ^c | 4.0 | 0.8 | 1.6–6.8 | –0.26 | 1.32 |
| IMI ^d | 4.9 | 1.0 | 1.3–7.0 | –0.61 | 0.44 |
| ART-fiction ^e | 6.3 | 6.0 | –1–39 | 1.84 | 4.46 |

Note. Unless otherwise noted, $N = 210$; RMET = Reading the Mind in the Eyes Task; EIT = Emotion Identification Task; MJ-Acc = Moral Judgment Task, Accidental Harm; MJ-Att = Moral Judgment Task, Attempted Harm; IRI-PT = Interpersonal Reactivity Index–Perspective-Taking, IRI-EC = Interpersonal Reactivity Index–Empathic Concern; IRI-FS = Interpersonal Reactivity Index–Fantasy Scale; BEES = Balanced Emotional Empathy Scale; SNI-P = Social Network Index–Number of People; SNI-ND = Social Network Index–Network Diversity; ULS = UCLA Loneliness Scale; MOS-SSS = Medical Outcomes Study–Social Support Scale; NTS = Narrative Transportation Scale (M across the 4 weeks of reading for fiction and nonfiction readers); IMI = Intrinsic Motivation Inventory; ART-fiction = Author Recognition Test–Fiction.

^a Values are reversed-scored such that higher scores indicate *less* moral permissibility (i.e., greater consideration of a character’s intention). ^b Values are reversed-scored such that higher scores indicate *less* loneliness. ^c $N = 130$. ^d $N = 132$. ^e $N = 207$.

Table 3
Zero-Order Correlations and Factors Loadings From the Measurement Model

| Measure | Zero-order correlations | | | | Factor loadings |
|------------------------|-------------------------|---------------------|---------------------|---|---------------------|
| Theory of mind | | | | | |
| | 1 | 2 | 3 | 4 | Theory of mind |
| 1. RMET | — | | | | .66 [0.45, 0.87]*** |
| 2. EIT | .39 [0.22, 0.57]*** | — | | | .57 [0.39, 0.74]*** |
| 3. MJ-Acc | .28 [0.15, 0.42]*** | .25 [0.12, 0.37]*** | — | | .46 [0.29, 0.63]*** |
| 4. MJ-Att ^a | .29 [0.14, 0.45]*** | .22 [0.07, 0.38]** | .27 [0.12, 0.39]*** | — | .45 [0.29, 0.61]*** |
| Empathy | | | | | |
| | 1 | 2 | 3 | 4 | Empathy |
| 1. IRI-EC | — | | | | .96 [0.86, 1.05]*** |
| 2. BEES | .77 [0.70, 0.83]*** | — | | | .81 [0.71, 0.91]*** |
| 3. IRI-FS | .41 [0.27, 0.53]*** | .39 [0.26, 0.50]*** | — | | .44 [0.29, 0.59]*** |
| 4. IRI-PT | .37 [0.24, 0.49]*** | .26 [0.12, 0.38]*** | .08 [−0.06, 0.21] | — | .37 [0.24, 0.50]*** |
| Social functioning | | | | | |
| | 1 | 2 | 3 | 4 | Social functioning |
| 1. ULS ^b | — | | | | .92 [0.72, 1.11]*** |
| 2. MOS-SSS | .73 [0.60, 0.80]*** | — | | | .79 [0.64, 0.95]*** |
| 3. SNI-P | .35 [0.22, 0.47]*** | .28 [0.16, 0.40]*** | — | | .38 [0.26, 0.50]*** |
| 4. SNI-ND | .33 [0.21, 0.44]*** | .26 [0.14, 0.37]*** | .91 [0.89, 0.93]*** | — | .35 [0.23, 0.47]*** |

Note. $N = 210$. Values under zero-order correlations represent Pearson r and 95% bias-corrected-and-accelerated CI [lower, upper] derived from 10,000 bootstrap samples; values under factor loadings represent standardized factors loadings and 95% CI from the measurement model CFA. RMET = Reading the Mind in the Eyes Task; EIT = Emotion Identification Task; MJ-Acc = Moral Judgment Task, Accidental Harm; MJ-Att = Moral Judgment Task, Attempted Harm; IRI-PT = Interpersonal Reactivity Index–Perspective-Taking; IRI-EC = Interpersonal Reactivity Index–Empathic Concern; IRI-FS = Interpersonal Reactivity Index–Fantasy Scale; BEES = Balanced Emotional Empathy Scale; SNI-P = Social Network Index–Number of People; SNI-ND = Social Network Index–Network Diversity; ULS = UCLA Loneliness Scale; MOS-SSS = Medical Outcomes Study–Social Support Scale; CI = confidence interval; CFA = confirmatory factor analysis.

^a Values are reversed-scored such that higher scores indicate *less* moral permissibility (i.e., greater consideration of a character's intention). ^b Values are reversed-scored such that higher scores indicate *less* loneliness.

* $p < .05$. ** $p < .01$. *** $p < .001$.

test $p < .05$), and the nonfiction group reported higher scores than the fiction group on BEES ($d = 0.42$, 95% CI [0.08, 0.76], Tukey HSD test $p < .05$). There were no other between-group differences on individual social measures.

Associations Between Fiction Reading and the Social Outcomes

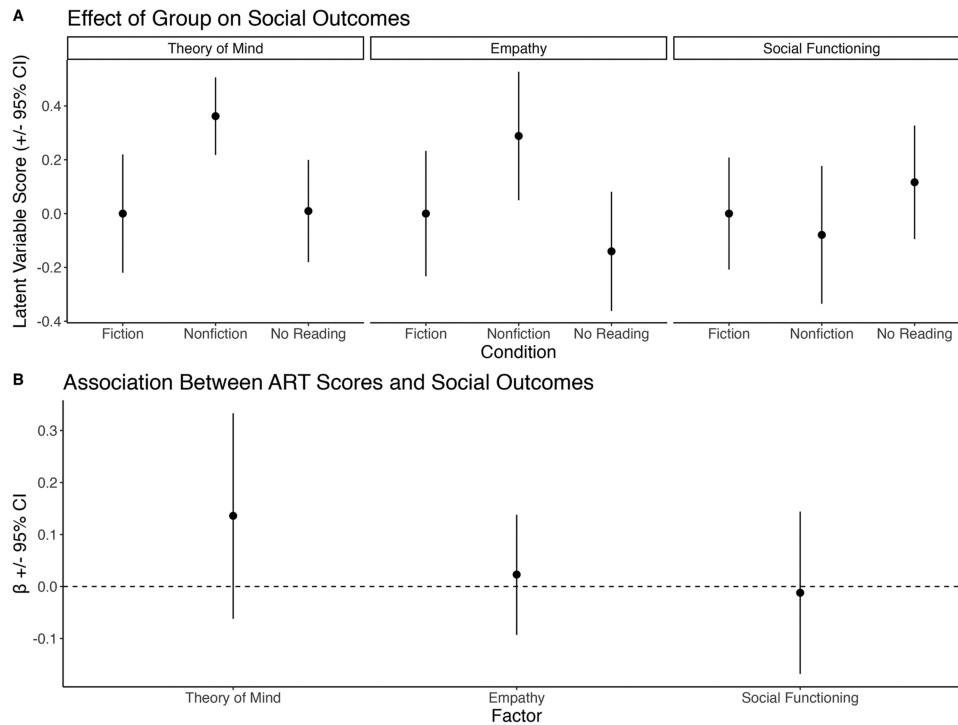
We first asked whether an intensive 4-week fiction reading manipulation was associated with better performance on a range of social outcomes postreading compared to an intensive 4-week nonfiction reading and an abstention from pleasure reading. To test this question, we conducted a SEM to evaluate the association between group assignment and the three latent variables determined in the aforementioned CFA. The SEM showed good model fit: $\chi^2(68) = 76.72$, $p = .219$, CFI = 0.99, RMSEA = 0.02 (90% CI [0.00, 0.05]), SRMR = 0.05. Compared to the no reading group, the nonfiction group demonstrated higher scores on the empathy latent variable (Figure 1A; Table 4). There were no other group differences. We reran this model controlling for demographic variables and WRAT score ($N = 170$). Model fit was adequate with absolute model fit indices (RMSEA, SRMR) indicating acceptable model fit— $\chi^2(104) = 155.17$, $p = .001$, CFI = 0.94, RMSEA = 0.05 (90% CI [0.04, 0.07]), SRMR = 0.06. The structural paths were similar, with the nonfiction group outperforming the no reading group, $\beta = -0.25$, $p = .010$, although at slightly greater magnitude.

Next, because lifetime fiction exposure is associated with social outcomes (Mumper & Gerrig, 2017), we asked whether the intensive fiction reading experience impacted performance after controlling for lifetime

fiction exposure. This model also allowed us to test the lifetime fiction exposure and social outcome association observed in other studies. To address this question, we conducted a similar SEM described above that included ART scores. The SEM showed good model fit: $\chi^2(77) = 87.34$, $p = .197$, CFI = 0.99, RMSEA = 0.03 (90% CI [0.00, 0.05]), SRMR = 0.05. The nonfiction group again outperformed the no reading group on the empathy latent variable (Table 4). In contrast to other work, we found no associations between ART scores and any of the latent variables (Figure 1B). We reran the model controlling for demographic variables and WRAT scores ($N = 167$). Model fit was adequate with absolute model fit indices (RMSEA, SRMR), indicating acceptable model fit— $\chi^2(113) = 163.07$, $p = .001$, CFI = 0.94, RMSEA = 0.05 (90% CI [0.03, 0.07]), SRMR = 0.06. The structural paths were similar, with the nonfiction group outperforming the no reading group, $\beta = -.24$, $p = .014$, on the empathy latent variable, and there being no association between ART scores and any of the latent variables.

Finally, we evaluated whether, as in some other work (Kidd & Castano, 2019), the impact of the intensive fiction reading experience was influenced by lifetime fiction reading. To evaluate this question, we conducted a similar SEM as the one described above that also included a group by ART interaction term. The SEM showed adequate model fit: $\chi^2(95) = 105.65$, $p = .214$, CFI = 0.99, RMSEA = 0.02 (90% CI [0.00, 0.04]), SRMR = 0.05. We observed no interactions between group and lifetime fiction exposure in predicting any of the latent variables (Table 4). In the follow-up SEM in which we controlled for demographic variables and WRAT scores, model fit was adequate, with absolute model fit indices (RMSEA, SRMR) indicating acceptable model fit— $\chi^2(131) = 187.30$, $p = .001$, CFI = 0.93, RMSEA = 0.05 (90% CI [0.03, 0.07]), SRMR = 0.05. The

Figure 1
Latent Variable Scores by Group and Associations With the Author Recognition Test



Note. ART = Author Recognition Test; CI = confidence interval.

structural paths were unchanged; we observed no interactions between group and ART in predicting any factors.

Associations Between Reading Engagement and the Social Outcomes

We evaluated the possibility that fiction's impact on social outcomes depends, in part, on aspects of narrative engagement; specifically, transportation into the narrative and intrinsic motivation to read. We tested this idea by generating factor scores for the three latent variables from the CFA and using multiple linear regression. First, we tested whether reading engagement was related to the outcomes irrespective of reading material by separately regressing each social outcome on narrative transportation and then intrinsic motivation. As the critical test to evaluate whether fiction specifically confers social benefits only for those readers who demonstrate high levels of narrative/task engagement, we next regressed each social outcome on reader engagement variable, reading material (fiction, nonfiction), and their interaction.

Narrative transportation was not related to any social outcome ($ps > .08$), nor did it moderate the impact of reading material on any of the social outcomes (condition by NTS interaction term $ps > .34$; Table S4 in the online supplemental materials). In contrast, intrinsic motivation was positively associated with empathy, $\beta = .35$, 95% CI [0.18, 0.51], $p < .001$, although it did not moderate the impact of reading material on empathy, interaction term $\beta = .12$, [-0.20, 0.45], $p = .457$. Intrinsic motivation was not related to the other social outcomes ($ps > .23$), nor did it moderate the association

between reading material and the other social outcomes (condition by IMI interaction term $ps > .054$; Table S5 in the online supplemental materials).

Discussion

Does fiction improve social ability in a way that might help people to understand one another? There are strong theoretical reasons to believe that the answer might be "yes" (Black et al., 2021; Mar, 2018b; Mar & Oatley, 2008; Mumper & Gerrig, 2019; Oatley, 2016). There is also empirical support behind the idea that lifetime fiction reading is positively associated with certain social outcomes (Mumper & Gerrig, 2017), though support for the idea that fiction causally improves social outcomes is more tenuous (Dodell-Feder & Tamir, 2018; Quinlan et al., 2023). Here, we evaluated this latter possibility by conducting a longitudinal randomized controlled study in which participants either read fiction or nonfiction, or did not engage in pleasure reading for 4 weeks. Compared to the laboratory methods typically used to address this question (i.e., single, short fiction exposure; assessment of a limited set of social outcomes with the same set of measures used in other studies), our study design was more intensive, better approximated how readers may engage with fiction in a real-world setting, while maintaining some control over the reading, and assessed a far richer set of social outcomes. Despite strengthening the reading manipulation in several significant ways and offering many more possible social outcomes, our main hypothesis, that fiction readers would show better performance on the social outcomes postreading, was not supported.

Table 4
Structural Equation Modeling Results

| Model | Dependent variable | Independent variable ^a | β [95% CI] | <i>p</i> | |
|--|--|-----------------------------------|----------------------------|--------------------|------|
| Social outcomes ~ Group (<i>N</i> = 210) | ToM | F vs. NF | .16 [−0.02, 0.35] | .092 | |
| | | F vs. NR | .00 [−0.22, 0.23] | .968 | |
| | | NF vs. NR | −.17 [−0.34, 0.00] | .056 | |
| | Empathy | F vs. NF | .13 [−0.06, 0.32] | .177 | |
| | | F vs. NR | −.07 [−0.25, 0.12] | .485 | |
| | | NF vs. NR | −.20 [−0.37, −0.04] | .019 | |
| | Social functioning | F vs. NF | −.04 [−0.21, 0.13] | .676 | |
| | | F vs. NR | .06 [−0.11, 0.22] | .504 | |
| | | NF vs. NR | .09 [−0.09, 0.27] | .312 | |
| | Social outcomes ~ Group + ART (<i>N</i> = 207) | ToM | F vs. NF | .17 [−0.02, 0.36] | .090 |
| | | | F vs. NR | .02 [−0.20, 0.24] | .864 |
| | | | NF vs. NR | −.15 [−0.33, 0.02] | .082 |
| ART | | | .14 [−0.06, 0.33] | .185 | |
| Empathy | | F vs. NF | .14 [−0.05, 0.33] | .143 | |
| | | F vs. NR | −.05 [−0.25, 0.14] | .609 | |
| | | NF vs. NR | −.20 [−0.37, −0.03] | .028 | |
| | | ART | .02 [−0.09, 0.14] | .702 | |
| Social functioning | | F vs. NF | −.04 [−0.21, 0.13] | .659 | |
| | | F vs. NR | .04 [−0.12, 0.21] | .607 | |
| | | NF vs. NR | .08 [−0.10, 0.27] | .381 | |
| | | ART | −.01 [−0.17, 0.14] | .884 | |
| Social outcomes ~ Group × ART (<i>N</i> = 207) | | ToM | F vs. NF × ART | −.03 [−0.31, 0.26] | .859 |
| | | | F vs. NR × ART | .03 [−0.24, 0.29] | .854 |
| | | | NF vs. NR × ART | .05 [−0.13, 0.23] | .614 |
| | Empathy | F vs. NF × ART | −.11 [−0.26, 0.05] | .178 | |
| | | F vs. NR × ART | −.07 [−0.25, 0.11] | .436 | |
| | | NF vs. NR × ART | .02 [−0.14, 0.17] | .839 | |
| | Social functioning | F vs. NF × ART | −.04 [−0.23, 0.14] | .644 | |
| | | F vs. NR × ART | −.15 [−0.37, 0.07] | .188 | |
| | | NF vs. NR × ART | −.11 [−0.30, 0.07] | .239 | |

Note. CI = confidence interval; F = fiction; NF = nonfiction; NR = no reading; ART = Author Recognition Test; ToM = theory of mind.

^a The first group listed in the independent variable column is the reference group. Bold values indicate $p < .05$.

Fiction readers outperformed neither nonfiction readers nor reading abstainers on any social outcome measure. Instead, nonfiction readers scored higher than people who abstained from pleasure reading on empathy. We also failed to replicate findings demonstrating an association between lifetime fiction exposure and social outcomes, and an interaction effect of lifetime fiction exposure and short fiction exposure on social outcomes. Why did we not observe an impact of fiction reading on the social outcomes despite using many of the same readings and outcome measures included in other work with a more intensive, ecologically valid reading experience? We see several strong possibilities, many of which speak to the limitations of our study design. We consider each of these in turn below.

First, although participants largely reported not reading nonstudy leisure material, it is possible that a small amount of this other material was enough to influence performance. Furthermore, while we asked

all participants to refrain from pleasure reading outside of the provided materials, participants were not able to abstain from work or school-related reading, which very well may have included fiction or social nonfiction (e.g., memoir, biography) for the nonfiction group (e.g., for participants studying English or history for example). We had no control over the other types of material consumed by participants during the study period. Research suggests that other narrative media such as television (Black & Barnes, 2015a) and video games (Bormann & Greitemeyer, 2016), positively impacts social outcomes. And, beyond consuming stories, producing stories in the form of writing—specifically when those stories are focused on people rather than whether those stories are fictional—also positively impacts social outcomes (Black & Barnes, 2021). It is possible that our participants may have been consuming or producing other such narratives during the four weeks in a way that may have influenced performance. To the

extent that participants randomly assigned to read nonfiction wanted to read fiction during the study period, they may have sought out these other forms of narrative fiction (e.g., television, podcasts) or person-focused material (e.g., biography).

Second, fiction's positive impact on social cognition may take two forms: an immediate and transient postreading prime and a slow-to-develop and durable social cognitive change. On the first account, fiction's social content and the social processes they invite may make the reader temporarily more interested and attentive to mental and emotional states, and temporarily activate the processes used to understand those states for use in other contexts (Lenhart & Richter, 2022; Mumper & Gerrig, 2019; Panero et al., 2016). This account is consistent with existing experimental work where participants are tested in the same laboratory session immediately after reading (cf., Lenhart & Richter, 2022). Since the final testing phase occurred within several days after the 4-week reading period, but not immediately after consuming any particular reading material, we would have missed the effect if it is, in fact, due to priming.

On the second account, fiction may lead to a durable positive change in social outcomes only after repeated engagement with fiction in which social cognitive processes may be practiced and honed. Repeated exposure to different social agents, interactions, and contexts in different texts allows for the opportunity to build a body of new social knowledge large enough to be used beneficially in real-world social interaction (Mar, 2018b). As part of this ongoing process, fiction may need an incubation period of sorts where the reader has time to evaluate, elaborate, and self-reflect on the material for it to have an impact on how they think about others (Appel & Richter, 2007; Bal et al., 2013). People may also simply need more time to accumulate fiction reading experience than the four weeks offered here. Both of these accounts are supported by the robust association between lifetime fiction reading and social outcomes (Mumper & Gerrig, 2017), and studies showing that fiction's positive social effect is observed not immediately after, but only after some time has passed (Appel & Richter, 2007; Bal et al., 2013). If this account is accurate, the final testing session may not have been long out enough after the end of the reading period for the necessary incubation period, and/or the amount of reading may not have entailed the necessary amount or intensity of social cognitive practice and social knowledge building to detect the longer-lasting changes. Another possibility related to these ideas is that reading multiple short stories, as opposed to full books, may compromise these processes, preventing readers from more fully engaging with the characters and narratives in the way that promotes social cognitive change. In potential support of this idea, some of the largest experimental effects of fiction reading comes from a study in which participants read entire books (Pino & Mazza, 2016).

Third, as demonstrated by others, fiction's social impact may depend on characteristics of the reader (Bal et al., 2013; Koopman, 2015, 2016; Schwerin & Lenhart, 2022; Tamir et al., 2016) and reading (Fong et al., 2013; Koopman, 2016; Schwerin et al., 2021). On this possibility, although we found an association between intrinsic motivation to read and empathy, this association was not specific to fiction readers. Said otherwise, participants who reported greater interest in and enjoyment of the readings, effort expended reading, increased choice in reading, and increased perceived value of the reading, whether it was fictional or nonfictional, were more likely to self-report higher levels of empathy. This finding converges with other studies demonstrating a positive association between intrinsic motivation and empathy (Findyartini et al., 2020; Oh & Roh, 2022), which

might simply speak to the idea that empathy is a motivated phenomenon (Zaki, 2014) that may be higher in people who generally demonstrate an approach orientation to tasks. Nevertheless, we only tested for two such characteristics of the reader, leaving open the possibility that other intrapersonal factors and text variables that we did not evaluate here strongly influenced whether fiction reading had a positive social impact.

Fourth, our study deviates from other empirical work by evaluating fiction's impact on latent social variables in which we incorporated a range of social indicators across a range of social constructs. We propose that this is a strength of the current approach, in that it allowed us to evaluate whether fiction's positive impact was on a class of variables versus a single measure. This issue is particularly important given that most studies have relied on a single measure, the RMET, which suffers from critical limitations (Baker et al., 2014; Black, 2019; Dodell-Feder et al., 2020; Peterson & Miller, 2012). To illustrate this point, in our dataset, we observe a positive association between RMET performance and ART scores (square-root-transformed as in other studies; Kidd & Castano, 2013; Panero et al., 2016; Samur et al., 2018), $r(205) = .14$, 95% CI $[-0.01, 0.25]$, $p = .043$, which is consistent with meta-analytic findings (Mumper & Gerrig, 2017), although the magnitude of association here is smaller. However, we did not observe an association between ART scores and the broader ToM latent variable, which included additional measures of mental state decoding and judging intentions when assigning blame for actions. This could suggest that, as a rule, fiction's positive impacts do not generalize to the broader category of a given social outcome. This may help to explain why we also did not observe associations between ART score and the social outcomes in the SEMs.

Finally, we note that several methodological issues may have contributed to the findings observed here. First, we may have been underpowered to detect effects including those tested in the SEMs and tests of the interaction between condition and transportation/intrinsic motivation in predicting outcomes. Second, RMET and ART scores were skewed, and demonstrated some range restriction. Although this is commonly observed for both measures (e.g., Black, 2019; Mar et al., 2006), this may have contributed to attenuated associations and false negatives. Third, although typical of RMET and the version of the BEES used here (Mehrabian, 2000; van Kuijk et al., 2018), both measures exhibited low internal consistency.

In summary, findings from our study suggest the following possibilities regarding fiction's social effects. First, contrary to other work, an intensive, several weeks long fiction reading experience in a context that better resembles real-world reading behavior does not positively impact social outcomes. That said, this does not preclude the existence of a priming effect or a positive effect of fiction reading after more prolonged exposures. This also does not preclude an effect of fiction reading on social outcomes that is smaller in magnitude than what we were able to detect with our sample size and the specific analysis approach we took here. Second, though lifetime fiction exposure is positively associated with certain measures of social cognition, these positive associations may not generalize to the broader social construct. Third, fiction reading may contribute to improved social ability, but only for certain readers who exhibit specific types of narrative engagement not tested here. Due to the limitations of this study, which include a modest sample size, issues in measurement, and limited control of nonstudy-related narrative consumption and production during the study period, we cannot conclusively rule-in or rule-out any of these possibilities. Nevertheless, this study provides no new evidence of a

beneficial effect of fiction reading on social outcomes. We recommend that in future work, researchers carefully delineate possible priming effects versus more durable changes with more naturalistic experiments that capitalize on positive aspects of reader engagement (e.g., transportation, intrinsic motivation to read; Mar, 2018b), and test for other characteristics of readers, readings, and their possible interactions (Koopman, 2018) that may influence fiction's social effects.

References

- Appel, M., & Richter, T. (2007). Persuasive effects of fictional narratives increase over time. *Media Psychology, 10*(1), 113–134. <https://doi.org/10.1080/15213260701301194>
- Arioli, M., Cattaneo, Z., Ricciardi, E., & Canessa, N. (2021). Overlapping and specific neural correlates for empathizing, affective mentalizing, and cognitive mentalizing: A coordinate-based meta-analytic study. *Human Brain Mapping, 42*(14), 4777–4804. <https://doi.org/10.1002/hbm.25570>
- Baker, C. A., Peterson, E., Pulos, S., & Kirkland, R. A. (2014). Eyes and IQ: A meta-analysis of the relationship between intelligence and “Reading the Mind in the Eyes”. *Intelligence, 44*, 78–92. <https://doi.org/10.1016/j.intell.2014.03.001>
- Bal, P. M., Veltkamp, M., & Young, L. (2013). How does fiction reading influence empathy? An experimental investigation on the role of emotional transportation. *PLoS ONE, 8*(1), Article e55341. <https://doi.org/10.1371/journal.pone.0055341>
- 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*(2), 241–251. <https://doi.org/10.1111/1469-7610.00715>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Black, J. E. (2019). An IRT analysis of the reading the mind in the eyes test. *Journal of Personality Assessment, 101*(4), 425–433. <https://doi.org/10.1080/00223891.2018.1447946>
- Black, J. E., & Barnes, J. L. (2015a). Fiction and social cognition: The effect of viewing award-winning television dramas on theory of mind. *Psychology of Aesthetics, Creativity, and the Arts, 9*(4), 423–429. <https://doi.org/10.1037/aca0000031>
- Black, J. E., & Barnes, J. L. (2015b). The effects of reading material on social and non-social cognition. *Poetics, 52*, 32–43. <https://doi.org/10.1016/j.poetic.2015.07.001>
- Black, J. E., & Barnes, J. L. (2021). Does writing promote social cognition? The role of fictionality and social content. *Psychology of Aesthetics, Creativity, and the Arts*. Advance online publication. <https://doi.org/10.1037/aca0000438>
- Black, J. E., Barnes, J. L., Oatley, K., Tamir, D. I., Dodell-Feder, D., Richter, T., & Mar, R. A. (2021). Stories and their role in social cognition. In D. Kuiken, & A. M. Jacobs (Eds.), *Handbook of empirical literary studies* (pp. 229–250). De Gruyter. <https://doi.org/10.1515/9783110645958-010>
- Bormann, D., & Greitemeyer, T. (2015). Immersed in virtual worlds and minds: Effects of in-game storytelling on immersion, need satisfaction, and affective theory of mind. *Social Psychological and Personality Science, 6*(6), 646–652. <https://doi.org/10.1177/1948550615578177>
- Boxell, L., Gentzkow, M., & Shapiro, J. (2020). *Cross-country trends in affective polarization* (Working Paper No. 26669). National Bureau of Economic Research. <https://doi.org/10.3386/w26669>
- Choi, J., Mogami, T., & Medalia, A. (2010). Intrinsic motivation inventory: An adapted measure for schizophrenia research. *Schizophrenia Bulletin, 36*(5), 966–976. <https://doi.org/10.1093/schbul/sbp030>
- Chopik, W. J., O'Brien, E., & Konrath, S. H. (2017). Differences in empathic concern and perspective taking across 63 countries. *Journal of Cross-Cultural Psychology, 48*(1), 23–38. <https://doi.org/10.1177/0022022116673910>
- Cohen, S. (1997). Social ties and susceptibility to the common cold. *JAMA, 277*(24), Article 1940. <https://doi.org/10.1001/jama.1997.03540480040036>
- Cotter, J., Granger, K., Backx, R., Hobbs, M., Looi, C. Y., & Barnett, J. H. (2018). Social cognitive dysfunction as a clinical marker: A systematic review of meta-analyses across 30 clinical conditions. *Neuroscience and Biobehavioral Reviews, 84*, 92–99. <https://doi.org/10.1016/j.neubiorev.2017.11.014>
- Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology, 44*(1), 113–126. <https://doi.org/10.1037/0022-3514.44.1.113>
- Djicic, M., Oatley, K., & Moldoveanu, M. C. (2013). Reading other minds: Effects of literature on empathy. *Scientific Study of Literature, 3*(1), 28–47. <https://doi.org/10.1075/ssol.3.1.06dji>
- Dodell-Feder, D., Ressler, K. J., & Germine, L. T. (2020). Social cognition or social class and culture? On the interpretation of differences in social cognitive performance. *Psychological Medicine, 50*(1), 133–145. <https://doi.org/10.1017/S003329171800404X>
- Dodell-Feder, D., & Tamir, D. I. (2018). Fiction reading has a small positive impact on social cognition: A meta-analysis. *Journal of Experimental Psychology: General, 147*(11), 1713–1727. <https://doi.org/10.1037/xge0000395>
- Findyartini, A., Felaza, E., Setyorini, D., & Mustika, R. (2020). Relationship between empathy and motivation in undergraduate medical students. *GMS Journal for Medical Education, 37*(4), Article Doc43. <https://doi.org/10.3205/ZMA001336>
- Finkel, E. J., Bail, C. A., Cikara, M., Ditto, P. H., Iyengar, S., Klar, S., Mason, L., McGrath, M. C., Nyhan, B., Rand, D. G., Skitka, L. J., Tucker, J. A., Van Bavel, J. J., Wang, C. S., & Druckman, J. N. (2020). Political sectarianism in America. *Science, 370*(6516), 533–536. <https://doi.org/10.1126/science.abe1715>
- Fong, K., Mullin, J. B., & Mar, R. A. (2013). What you read matters: The role of fiction genre in predicting interpersonal sensitivity. *Psychology of Aesthetics, Creativity, and the Arts, 7*(4), 370–376. <https://doi.org/10.1037/a0034084>
- Germine, L., & Hooker, C. I. (2011). Face emotion recognition is related to individual differences in psychosis-proneness. *Psychological Medicine, 41*(5), 937–947. <https://doi.org/10.1017/S0033291710001571>
- Gerrig, R. J. (1993). *Experiencing narrative worlds: On the psychological activities of reading*. Yale University Press.
- Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences, 102*, 74–78. <https://doi.org/10.1016/j.paid.2016.06.069>
- Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology, 79*(5), 701–721. <https://doi.org/10.1037/0022-3514.79.5.701>
- Green, M. F., Horan, W. P., & Lee, J. (2015). Social cognition in schizophrenia. *Nature Reviews Neuroscience, 16*(10), 620–631. <https://doi.org/10.1038/nrn4005>
- Hall, J., & Leary, M. (2020, September 17). The U.S. has an empathy deficit. *Scientific American*. <https://www.scientificamerican.com/article/the-us-has-an-empathy-deficit/>
- Harrell, F. E. (2022). *Hmisc: Harrell miscellaneous* (Version 4.7-0) [Computer software]. <https://CRAN.R-project.org/package=Hmisc>
- Hartung, F., & Willems, R. M. (2020). *Amount of fiction reading correlates with higher connectivity between cortical areas for language and mentalizing* [Preprint]. Scientific Communication and Education. <https://doi.org/10.1101/2020.06.08.139923>
- Hogan, P. C. (2003). *The mind and its stories: Narrative universals and human emotion* (1st ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511499951>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. <https://doi.org/10.1080/1070519990540118>
- Jobst, L. J., Bader, M., & Moshagen, M. (2023). A tutorial on assessing statistical power and determining sample size for structural equation models. *Psychological Methods, 28*(1), 207–221. <https://doi.org/10.1037/met0000423>

- Johnson, D. R. (2012). Transportation into a story increases empathy, prosocial behavior, and perceptual bias toward fearful expressions. *Personality and Individual Differences, 52*(2), 150–155. <https://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*(3), 306–312. <https://doi.org/10.1037/a0033261>
- Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2018). *semTools: Useful tools for structural equation modeling* (R package Version 0.5-3) [Computer software]. <https://CRAN.R-project.org/package=semTools>
- Kassambara, A. (2021). *rstatix: Pipe-Friendly Framework for Basic Statistical Tests (Version 0.7.0)* [Computer software]. <https://CRAN.R-project.org/package=rstatix>
- Ken, K. (2022). *MBESS: The MBESS R Package (Version 4.9.2)* [Computer software]. <https://CRAN.R-project.org/package=MBESS>
- Kidd, D., & Castano, E. (2013). Reading literary fiction improves theory of mind. *Science, 342*(6156), 377–380. <https://doi.org/10.1126/science.1239918>
- Kidd, D., & Castano, E. (2019). Reading literary fiction and theory of mind: Three preregistered replications and extensions of Kidd and Castano (2013). *Social Psychological and Personality Science, 10*(4), 522–531. <https://doi.org/10.1177/1948550618775410>
- Kidd, D., Ongis, M., & Castano, E. (2016). On literary fiction and its effects on theory of mind. *Scientific Study of Literature, 6*(1), 42–58. <https://doi.org/10.1075/ssol.6.1.04kid>
- Kittel, A. F. D., Olderbak, S., & Wilhelm, O. (2022). Sty in the mind's eye: A meta-analytic investigation of the nomological network and internal consistency of the “reading the mind in the eyes” test. *Assessment, 29*(5), 872–895. <https://doi.org/10.1177/1073191121996469>
- Konrath, S. H., O'Brien, E. H., & Hsing, C. (2011). Changes in dispositional empathy in American college students over time: A meta-analysis. *Personality and Social Psychology Review, 15*(2), 180–198. <https://doi.org/10.1177/1088868310377395>
- Koopman, E. M. (2015). Empathic reactions after reading: The role of genre, personal factors and affective responses. *Poetics, 50*, 62–79. <https://doi.org/10.1016/j.poetic.2015.02.008>
- Koopman, E. M. (2016). Effects of “literariness” on emotions and on empathy and reflection after reading. *Psychology of Aesthetics, Creativity, and the Arts, 10*(1), 82–98. <https://doi.org/10.1037/aca0000041>
- Koopman, E. M. (2018). Does originality evoke understanding? The relation between literary reading and empathy. *Review of General Psychology, 22*(2), 169–177. <https://doi.org/10.1037/gpr0000107>
- Kou, X., Konrath, S., & Goldstein, T. R. (2020). The relationship among different types of arts engagement, empathy, and prosocial behavior. *Psychology of Aesthetics, Creativity, and the Arts, 14*(4), 481–492. <https://doi.org/10.1037/aca0000269>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). LmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1–26. <https://doi.org/10.18637/jss.v082.i13>
- Kyriazos, T. A. (2018). Applied psychometrics: Sample size and sample power considerations in factor analysis (EFA, CFA) and SEM in general. *Psychology, 09*(08), 2207–2230. <https://doi.org/10.4236/psych.2018.98126>
- Lecce, S., Bianco, F., & Hughes, C. (2021). Reading minds and reading texts: Evidence for independent and specific associations. *Cognitive Development, 57*, Article 101010. <https://doi.org/10.1016/j.cogdev.2021.101010>
- Lenhart, J., & Richter, T. (2022). Does reading a single short story of literary fiction improve social-cognitive skills? Testing the priming hypothesis. *Psychology of Aesthetics, Creativity, and the Arts*. Advance online publication. <https://doi.org/10.1037/aca0000533>
- Lenth, R. V. (2022). *emmeans: Estimated marginal means, aka least-squares means (Version 1.7.4-1)* [Computer software]. <https://CRAN.R-project.org/package=emmeans>
- Lüdtke, D., Ben-Shachar, M., Patil, I., Waggoner, P., & Makowski, D. (2021). Performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software, 6*(60), Article 3139. <https://doi.org/10.21105/joss.03139>
- Mak, H. W., & Fancourt, D. (2020). Longitudinal associations between reading for pleasure and child maladjustment: Results from a propensity score matching analysis. *Social Science and Medicine, 253*, Article 112971. <https://doi.org/10.1016/j.socscimed.2020.112971>
- Mar, R. A. (2004). The neuropsychology of narrative: Story comprehension, story production and their interrelation. *Neuropsychologia, 42*(10), 1414–1434. <https://doi.org/10.1016/j.neuropsychologia.2003.12.016>
- Mar, R. A. (2011). The neural bases of social cognition and story comprehension. *Annual Review of Psychology, 62*(1), 103–134. <https://doi.org/10.1146/annurev-psych-120709-145406>
- Mar, R. A. (2018a). Stories and the promotion of social cognition. *Current Directions in Psychological Science, 27*(4), 257–262. <https://doi.org/10.1177/0963721417749654>
- Mar, R. A. (2018b). Evaluating whether stories can promote social cognition: Introducing the social processes and content entrained by narrative (SPaCEN) framework. *Discourse Processes, 55*(5–6), 454–479. <https://doi.org/10.1080/0163853X.2018.1448209>
- Mar, R. A., & Oatley, K. (2008). The function of fiction is the abstraction and simulation of social experience. *Perspectives on Psychological Science, 3*(3), 173–192. <https://doi.org/10.1111/j.1745-6924.2008.00073.x>
- Mar, R. A., Oatley, K., Hirsh, J., dela Paz, J., & Peterson, J. B. (2006). Bookworms versus nerds: Exposure to fiction versus non-fiction, divergent associations with social ability, and the simulation of fictional social worlds. *Journal of Research in Personality, 40*(5), 694–712. <https://doi.org/10.1016/j.jrp.2005.08.002>
- Mar, R. A., Oatley, K., & Peterson, J. B. (2009). Exploring the link between reading fiction and empathy: Ruling out individual differences and examining outcomes. *Communications, 34*(4), 407–428. <https://doi.org/10.1515/COMM.2009.025>
- Mason, R. A., & Just, M. A. (2009). The role of the theory-of-mind cortical network in the comprehension of narratives: Theory of mind in narrative processing. *Language and Linguistics Compass, 3*(1), 157–174. <https://doi.org/10.1111/j.1749-818X.2008.00122.x>
- Mayer, M. (2022). *confint: Confidence Intervals* (R package Version 0.1.2) [Computer software]. <https://CRAN.R-project.org/package=confint>
- Mehrabian, A. (2000). Beyond IQ: Broad-based measurement of individual success potential or “emotional intelligence”. *Genetic, Social, and General Psychology Monographs, 126*(2), 133–239.
- Mehrabian, A., & Epstein, N. (1972). A measure of emotional empathy. *Journal of Personality, 40*(4), 525–543. <https://doi.org/10.1111/j.1467-6494.1972.tb00078.x>
- Moran, J. M., Young, L. L., Saxe, R., Lee, S. M., O'Young, D., Mavros, P. L., & Gabrieli, J. D. (2011). Impaired theory of mind for moral judgment in high-functioning autism. *Proceedings of the National Academy of Sciences, 108*(7), 2688–2692. <https://doi.org/10.1073/pnas.1011734108>
- Mumper, M. L., & Gerrig, R. J. (2017). Leisure reading and social cognition: A meta-analysis. *Psychology of Aesthetics, Creativity, and the Arts, 11*(1), 109–120. <https://doi.org/10.1037/aca0000089>
- Mumper, M. L., & Gerrig, R. J. (2019). How does leisure reading affect social cognitive abilities? *Poetics Today, 40*(3), 453–473. <https://doi.org/10.1215/03335372-7558080>
- Nijhof, A. D., Willems, R. M., & Kotz, S. (2015). Simulating fiction: Individual differences in literature comprehension revealed with fMRI. *PLoS ONE, 10*(2), Article e0116492. <https://doi.org/10.1371/journal.pone.0116492>
- Oatley, K. (1999). Why fiction may be twice as true as fact: Fiction as cognitive and emotional simulation. *Review of General Psychology, 3*(2), 101–117. <https://doi.org/10.1037/1089-2680.3.2.101>
- Oatley, K. (2016). Fiction: Simulation of social worlds. *Trends in Cognitive Sciences, 20*(8), 618–628. <https://doi.org/10.1016/j.tics.2016.06.002>

- Ochsner, K. N. (2008). The social-emotional processing stream: Five core constructs and their translational potential for schizophrenia and beyond. *Biological Psychiatry*, *64*(1), 48–61. <https://doi.org/10.1016/j.biopsych.2008.04.024>
- Oh, S., & Roh, S.-C. (2022). Intrinsic motivation for work activities is associated with empathy: Investigating the indirect relationship between intrinsic motivation for work activities and social support through empathy and prosocial behavior. *Personality and Individual Differences*, *189*, Article 111487. <https://doi.org/10.1016/j.paid.2021.111487>
- Olderbak, S., Wilhelm, O., Oltar, G., Geiger, M., Brennenman, M. W., & Roberts, R. D. (2015). A psychometric analysis of the reading the mind in the eyes test: Toward a brief form for research and applied settings. *Frontiers in Psychology*, *6*, Article 1503. <https://doi.org/10.3389/fpsyg.2015.01503>
- 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. <https://doi.org/10.1037/pspa0000064>
- Perrin, Andrew. (2016). *Book Reading 2016*. <https://www.pewresearch.org/internet/2016/09/01/book-reading-2016/>
- Peterson, E., & Miller, S. F. (2012). The eyes test as a measure of individual differences: How much of the variance reflects verbal IQ? *Frontiers in Psychology*, *3*, Article 220. <https://doi.org/10.3389/fpsyg.2012.00220>
- Pino, M. C., & Mazza, M. (2016). The use of “literary fiction” to promote mentalizing ability. *PLoS ONE*, *11*(8), Article e0160254. <https://doi.org/10.1371/journal.pone.0160254>
- Quinlan, J. A., Padgett, J. K., Khajehnasiri, A., & Mar, R. A. (2023). Does a brief exposure to literary fiction improve social ability? Assessing the evidential value of published studies with a p-curve. *Journal of Experimental Psychology: General*, *152*(3), 723–732. <https://doi.org/10.1037/xge0001302>
- Rain, M., & Mar, R. A. (2014). Measuring reading behavior: Examining the predictive validity of print-exposure checklists. *Empirical Studies of the Arts*, *32*(1), 93–108. <https://doi.org/10.2190/EM.32.1f>
- R Core Team. (2022). *R: A language and environment for statistical computing* (Version 4.2.0) [Computer software]. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Revelle, W. (2018). *Psych: Procedures for personality and psychological research* (R package Version 2.2.5) [Computer software]. Northwestern University. <https://CRAN.R-project.org/package=psychVersion=1.8.12>
- Rossee, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. <https://doi.org/10.18637/jss.v048.i02>
- RStudio Team. (2020). *RStudio: Integrated development for R* (Version 1.3.1093) [Computer software]. <https://www.rstudio.com/>
- Russell, D. W. (1996). UCLA loneliness scale (version 3): Reliability, validity, and factor structure. *Journal of Personality Assessment*, *66*(1), 20–40. https://doi.org/10.1207/s15327752jpa6601_2
- Samur, D., Tops, M., & Koole, S. L. (2018). Does a single session of reading literary fiction prime enhanced mentalising performance? Four replication experiments of Kidd and Castano (2013). *Cognition and Emotion*, *32*(1), 130–144. <https://doi.org/10.1080/02699931.2017.1279591>
- Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality*, *47*(5), 609–612. <https://doi.org/10.1016/j.jrp.2013.05.009>
- Schurz, M., Radua, J., Tholen, M. G., Maliske, L., Margulies, D. S., Mars, R. B., Sallet, J., & Kanske, P. (2021). Toward a hierarchical model of social cognition: A neuroimaging meta-analysis and integrative review of empathy and theory of mind. *Psychological Bulletin*, *147*(3), 293–327. <https://doi.org/10.1037/bul0000303>
- Schwerin, J., & Lenhart, J. (2022). The effects of literariness on social-cognitive skills: Examining narrative engagement, transportation, and identification as moderators. *Psychology of Aesthetics, Creativity, and the Arts*. Advance online publication. <https://doi.org/10.1037/aca0000514>
- Schwering, S. C., Ghaffari-Nikou, N. M., Zhao, F., Niedenthal, P. M., & MacDonald, M. C. (2021). Exploring the relationship between fiction reading and emotion recognition. *Affective Science*, *2*(2), 178–186. <https://doi.org/10.1007/s42761-021-00034-0>
- Sherbourne, C. D., & Stewart, A. L. (1991). The MOS social support survey. *Social Science and Medicine*, *32*(6), 705–714. [https://doi.org/10.1016/0277-9536\(91\)90150-B](https://doi.org/10.1016/0277-9536(91)90150-B)
- 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*(2), 215–224. <https://doi.org/10.1093/scan/nsv114>
- The Rise of E-Reading*. (2012). <https://www.pewresearch.org/internet/2012/04/04/the-rise-of-e-reading-5/>
- van der Kleij, S. W., Apperly, I., Shapiro, L. R., Ricketts, J., & Devine, R. T. (2022). Reading fiction and reading minds in early adolescence: A longitudinal study. *Journal of Experimental Child Psychology*, *222*, Article 105476. <https://doi.org/10.1016/j.jecp.2022.105476>
- van Kuijk, I., Verkoijen, P., Dijkstra, K., Zwaan, R. A., & Vul, E. (2018). The effect of reading a short passage of literary fiction on theory of mind: A replication of Kidd and Castano (2013). *Collabra: Psychology*, *4*(1), Article 7. <https://doi.org/10.1525/collabra.117>
- Venables, W. N., Ripley, B. D., & Venables, W. N. (2002). *Modern applied statistics with S* (4th ed.). Springer.
- West, R. F., Stanovich, K. E., & Mitchell, H. R. (1993). Reading in the real world and its correlates. *Reading Research Quarterly*, *28*(1), Article 34. <https://doi.org/10.2307/747815>
- Wiley, J. F. (2020). *multilevelTools: Multilevel and mixed effects model diagnostics and effect sizes* (R package Version 0.1.1) [Computer software]. <https://CRAN.R-project.org/package=multilevelTools>
- Wilkinson, G. S., & Robertson, G. J. (2006). *Wide range achievement test 4 professional manual*. Psychological Assessment Resources.
- Willems, R. M., & Hartung, F. (2017). *Engaging regularly with fiction influences connectivity in cortical areas for language and mentalizing* [Preprint]. PsyArXiv. <https://doi.org/10.31234/osf.io/e7bqj>
- Willems, R. M., & Jacobs, A. M. (2016). Caring about dostoyevsky: The untapped potential of studying literature. *Trends in Cognitive Sciences*, *20*(4), 243–245. <https://doi.org/10.1016/j.tics.2015.12.009>
- Young, L., Camprodon, J. A., Hauser, M., Pascual-Leone, A., & Saxe, R. (2010). Disruption of the right temporoparietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments. *Proceedings of the National Academy of Sciences*, *107*(15), 6753–6758. <https://doi.org/10.1073/pnas.0914826107>
- Young, L., Cushman, F., Hauser, M., & Saxe, R. (2007). The neural basis of the interaction between theory of mind and moral judgment. *Proceedings of the National Academy of Sciences*, *104*(20), 8235–8240. <https://doi.org/10.1073/pnas.0701408104>
- Young, L., Koenigs, M., Kruepke, M., & Newman, J. P. (2012). Psychopathy increases perceived moral permissibility of accidents. *Journal of Abnormal Psychology*, *121*(3), 659–667. <https://doi.org/10.1037/a0027489>
- Zaki, J. (2014). Empathy: A motivated account. *Psychological Bulletin*, *140*(6), 1608–1647. <https://doi.org/10.1037/a0037679>
- Zaki, J., & Ochsner, K. (2012). The neuroscience of empathy: Progress, pitfalls and promise. *Nature Neuroscience*, *15*(5), 675–680. <https://doi.org/10.1038/nn.3085>
- Zhang, Z., & Mai, Y. (2023). *WebPower: Basic and advanced statistical power analysis* (Version 0.8.7) [Computer software]. <https://CRAN.R-project.org/package=WebPower>
- Zwaan, R. A. (2016). Situation models, mental simulations, and abstract concepts in discourse comprehension. *Psychonomic Bulletin and Review*, *23*(4), 1028–1034. <https://doi.org/10.3758/s13423-015-0864-x>

Received December 15, 2022

Revision received May 5, 2023

Accepted June 7, 2023 ■