Cognitive Effects and Correlates of Reading Fiction: Two Pre-Registered Multi-Level Meta-Analyses

Lena Wimmer[1], Gregory Currie[2], Stacie Friend[3], Jörg Wittwer[1], and Heather J Ferguson[4]

[1] Department of Education, University of Freiburg

[2] Department of Philosophy, University of York

[3] Department of Philosophy, Birkbeck, University of London

[4] School of Psychology, University of Kent

Author Note

Lena Wimmer https://orcid.org/0000-0001-8434-1766

Gregory Currie https://orcid.org/0000-0001-8364-997X

Stacie Friend https://orcid.org/0000-0001-7791-8467

Jörg Wittwer https://orcid.org/0000-0001-9984-7479

Heather J Ferguson https://orcid.org/0000-0002-1575-4820

Word count: 14,139 words

Preliminary versions of the data appearing in the manuscript were presented at conferences/meetings, but were not posted on a listserv, shared on a website, etc.

SF and HF were funded by a Templeton Religion Trust grant, grant number TRT0476.

We have no known conflict of interest to disclose.

These syntheses have been registered at the Open Science Framework, https://osf.io/mskcz/?view_only=c6d5a6dbde2a4d0c88dbfdf5a72e30ab

All meta-analytic data, analysis code, and research materials (including our coding scheme) are available at https://osf.io/dus2w/?view_only=bb8790263de94ccfa24eecf3be5c3833.

Correspondence concerning this article should be addressed to Lena Wimmer, University of Freiburg, Department of Education, Rempartstr. 11, D-79098 Freiburg im Breisgau, Germany. Email: lena.wimmer@ezw.uni-freiburg.de

Abstract

Despite significantly increasing research efforts, the psychological effects of reading fiction remain under debate. We present two pre-registered meta-analyses synthesizing cognitive effects and correlates of reading fiction. In meta-analysis 1 (371 effect sizes/70 experiments), reading fiction led to significant small-sized cognitive benefits, g = 0.14, 95% *CI* [0.06, 0.21]. This effect of fiction reading was moderated by comparison group (effects were greater when reading fiction was compared with watching fiction or reading nothing than when reading fiction was compared with reading nonfiction) and outcome variable (significant effects emerged for empathy and mentalizing only). In meta-analysis 2 (559 effect sizes/114 studies reporting correlations), lifetime exposure to print fiction was linked with significant small-sized cognitive benefits, r = .16, 95% *CI* [.13, .19]. This effect was moderated by outcome variable (effects were greatest for verbal abilities, followed by general cognitive abilities and empathy/mentalizing/outgroup judgments), fictionality of the print material (greater effects were found for fiction than nonfiction), publication status (published work exhibited greater effects than unpublished work), type of assessment measure (larger effects emerged when neither the outcome nor print exposure were assessed via self-report, than when either the outcome or print exposure were assessed via self-report), participant group (community samples showed greater effects than student samples), study design (greater effects were found for correlational than for experimental designs), and percentage of female participants (via a negative relationship with cognitive benefits). Together, these meta-analyses provide robust evidence for a small-sized positive relationship between reading fiction and cognitive benefits.

Keywords: fiction, cognition, empathy, theory of mind, meta-analysis

Public Significance Statement

This research project suggests that people who read a lot of fiction have better cognitive skills than people who read little or no fiction. These benefits are small in size across various cognitive skills, but of medium size for verbal and general cognitive abilities, for example, intelligence. Importantly, there is a stronger association between reading fiction and cognitive skills than between reading nonfiction and those skills. However, whether the benefits are caused by reading fiction or by one or more other variables remains to be determined through future research.

Cognitive Effects and Correlates of Reading Fiction: Two Pre-Registered Multi-Level Meta-Analyses

We regularly encounter written fiction in our everyday lives. Increasing rates of reading fiction during the Covid-19 pandemic (Anuar et al., 2021; Cho et al., 2021; Davies et al., 2022) illustrate the important role that fiction plays in humans' lives, even though attitudes toward reading (as indexed by how much students and their parents report liking reading) have become more negative internationally over the preceding decades (Hooper, 2020). While reading fiction is indisputably a popular activity, its psychological effects are debated (e.g., Best, 2021). The present article reports two meta-analyses synthesizing the cognitive effects of fiction-reading assignments using experiments and the cognitive correlates of lifetime exposure to written fiction.

Psychological Effects of Reading Fiction

Fictions, in the sense relevant here, can be conceived roughly as texts understood not to be intended as factually true, at least in large part; they typically involve reference to non-real events, objects, and persons (e.g., Gertken & Köppe, 2009). Contemporary pragmatic approaches (e.g., Currie, 1985; Eco, 1994; Lamarque & Olsen, 1994) stress that fictions are based on converging intentions of a work's author and its readers (see also the concept of the writer-reader contract; Tierney & LaZansky, 1980) which in turn evolve from cultural conventions. For example, the content of *The Count of Monte Christo* (Alexandre Dumas) was made up by its author, and this is typically recognized by readers. In sum, non-fictional works are generally expected to be true to reality, whereas this requirement is relaxed for fictional ones (Friend, 2012).

Although entertainment is widely identified as the immediate purpose of most fiction consumption (e.g., Dubourg & Baumard, 2022), written fictions are also widely promoted as educational tools (e.g., Erikson et al., 2020; Schultz, 2020; Vrasidas et al., 2015). So, it is

important for us to know what sorts of positive outcomes (if any) they tend to foster. In the next paragraphs, we provide an overview of some of the key issues in this controversy.

Empathy and Theory of Mind (ToM)

Regarding the effects of reading fiction, the field of social cognition (referring to the perception, interpretation, and response to social information; Fiske & Taylor, 2013) has attracted the most scholarly activity in recent years; special attention has been given to empathy and theory of mind. Generally speaking, *empathy* is an umbrella term for processes that help people share and comprehend others' affects (Ferguson & Wimmer, 2023; Håkansson Eklund & Summer Meranius, 2021; note that we consider this a working definition as there is at present no generally accepted definition of empathy; see also Happé et al., 2017). Empathy is often considered multi-faceted and hierarchical in that it involves lower- and higher-level processes (Schurz et al., 2021; Zurek & Scheithauer, 2017). Emotion recognition – decoding others' emotional states – exemplifies a lower-level process (Preston et al., 2020). Affective or emotional empathy, that is, sharing others' emotions while conscious that one's emotional response is triggered by that of the other person (Cuff et al., 2016), and cognitive empathy, that is, understanding others' emotional state (Preston et al., 2020), are higher-level processes. ToM, also known as mentalizing, is about comprehending mental states in general, not only emotional ones (Preston et al., 2020).

Some theoretical approaches (e.g., Mar, 2018a; Mumper & Gerrig, 2019) predict that reading stories – including but not exclusively fictional ones – improves empathy and ToM. As opposed to this, Mar and Oatley's (2008) simulation model assumes that reading fictional rather than nonfictional narratives especially has these benefits. More precisely, it maintains that fictional narratives, which are mostly about social relationships and interactions, function as simulations of the social world, encouraging readers to imagine fictitious situations which abstract from reality. The fictionality of the simulations allows readers to adopt protagonists' perspectives in a distanced way, while the abstraction and simplification enable them to grasp

the rules of the social world more clearly. Increases in empathy and ToM are predicted outcomes of these imagined simulations.

Indeed, several studies have suggested that reading short fictional texts enhances empathy and ToM (for a review, see Dodell-Feder & Tamir, 2018), and that lifetime exposure to print fiction is positively correlated with both capacities (for a review, see Mumper & Gerrig, 2017). However, two central issues remain unresolved. The first is the robustness of these findings in view of multiple failed replications and studies with null results (see, e.g., Chlebuch et al., 2020; Lenhart & Richter, 2022; Panero et al., 2016; Samur et al., 2018; Wimmer et al., 2021a, 2022a). Second, a capacity for high levels of empathy and ToM is generally considered desirable (e.g., Ferguson et al., 2021; Goldstein & Winner, 2012) and research confirming that reading fiction has robust positive effects on these capacities is likely to be welcomed. It is worth bearing in mind, however, the possibility that fiction has adverse effects in these areas. By adverse effects we mean either an increase in a generally undesirable type of cognition, for instance racist attitudes or stereotypical or inaccurate beliefs, or a decrease in an inherently desirable type of cognition, for instance prosocial or accurate beliefs. According to Best (2020), engaging the brain network that subserves empathy and ToM during reading fiction may lead to source monitoring errors; readers may not be able to link their knowledge with its source, so cannot reliably determine whether knowledge has been acquired from the real or from a fictional world. In that case, readers may seek to use schemas in the real world which are valid only in a fictional world; similarly, inferences drawn during reading fiction could be mistakenly transferred to the real world. It is also possible that empathizing with fictional characters may in some circumstances reduce a person's capacity for empathy (Currie, 2020, Section 11.7, drawing on work on "moral self-licensing" by, e.g., Effron & Conway, 2015).

Knowledge, Biases, and Prejudices

Similar disputes about whether the effects of fiction are positive or negative are familiar in two other domains. (It is important to note that there need not be uniform answers in any of these cases: effects may be positive for some fictions in some contexts, and negative in others.) First, it is often suggested that written fiction is a source of knowledge in the sense that it enables readers of all ages to acquire correct information and desirable attitudes (Best, 2021). Note however that a study by Wimmer et al. (2021a) found that frequent readers of fiction and frequent readers of nonfiction demonstrate equally high levels of general knowledge. Similarly, Hopkins and Weisberg (2017) in their review state that "the current body of evidence does not allow us to conclude definitively whether learning from fiction is the same, better, or worse than learning from other sources" (p. 62). What is more, a number of experiments have demonstrated that readers may easily acquire incorrect information from fiction even though they are aware that authors of fiction do not stick reliably to the truth (e.g., Appel & Richter, 2007; Prentice et al., 1997; Rapp et al., 2014; Wheeler et al., 1999). A question unresolved by current research is whether beliefs acquired from fiction are on balance more likely to be correct than incorrect.

Second, there is a dispute regarding fiction-based effects on biases and prejudices. A number of experiments have shown that reading fiction can be used effectively for bias and prejudice reduction, for example, in terms of race bias (Johnson et al., 2014), prejudice against Arab Muslims (Johnson et al., 2013), and attitudes towards stigmatized groups more broadly (Vezzali et al., 2015). However, according to Goffin and Friend (2022), although these experiments provide evidence that fictions can reduce harmful biases and prejudices, they fail to demonstrate that fictions do not also increase them. In fact, these authors give reasons to think that adverse effects are more likely than positive ones.

The earlier mentioned simulation model (Mar & Oatley, 2008) does not pertain to fiction-based effects on knowledge, biases, and prejudices, particularly if these cognitions are not social in content. A theoretical approach concerning both social and non-social cognitive

effects of reading fiction has been put forward by Consoli (2018). Whether reading fiction leads to desirable or undesirable cognitive outcomes is thought to depend on the accessibility of text content – highly accessible content would lead to shallow processing associated with unwanted outcomes, whereas less accessible content would lead to increased epistemic vigilance linked with cognitive benefits. The impact of accessibility on cognitive effects is conditioned on a number of variables including reader characteristics (e.g., motivation, experience with subject matter, or personality). In sum, there are issues about the types of psychological outcomes that are influenced by reading fiction and about the desirability of these outcomes. These and other continuing debates over the cognitive effects of reading fiction highlight the need for a systematic literature review, as provided in a quantitative way by meta-analyses, to help resolve those debates.

Previous Meta-Analyses

Two existing meta-analyses can be viewed as steps in this direction, with both focusing specifically on social cognitive outcomes. First, Dodell-Feder and Tamir (2018) synthesized 51 effect sizes from 14 experimental studies in which participants were assigned short reading tasks. Studies were included if they followed a true experimental design with random allocation to condition; if reading fiction was compared with either reading nothing or reading nonfiction; and if social cognition was assessed as a dependent measure. When compared with control conditions, reading fiction was found to have significant but small-sized benefits for social cognition, $g = 0.15$. Moderator analyses tested whether this overall effect was modulated by the following variables: sample type (i.e., whether research participants were students or recruited from Amazon's Mechanical Turk – this could be critical due to differing incentives [i.e., participation in exchange for course credits *vs.* monetary compensation]); type of comparison group (i.e., whether reading fiction was compared with reading nothing or reading nonfiction – this is relevant to the question whether effects are driven by reading fiction in particular or reading in general); type of dependent

measure (i.e., whether dependent measures were based on behavioral performance or on self-report – this is important since self-report measures have limited validity, especially for assessing cognition [e.g., Williams et al., 2017]); type of outcome (i.e., whether mentalizing or emotion sharing was assessed – given that these are considered related but distinguishable constructs, both outcomes could be differently affected by reading fiction); participant age (this could be of interest due to cross-generational alterations in reading habits and cognition's sensitivity to developmental changes); the percentage of female participants in the sample (this seems reasonable since females have a stronger preference for fiction than males [e.g., Hu et al., 2023; Summers, 2013] and also tend to read more than males overall [Bradshaw & Nichols, 2004]).

None of these variables was found to moderate the aggregate meta-analytic effect. To examine publication bias, the publication status of each study (i.e., whether it was published or not) was incorporated as a moderator variable, and a funnel plot and Egger's test were implemented. None of these tests suggested the presence of publication bias. Applying the same criteria for study inclusion as Dodell-Feder and Tamir (2018; except that only published work, written in English, was deemed eligible), Quinlan et al. (2022) investigated the evidential value of this body of research through a *p*-curve analysis. This analysis involved 13 studies, eight of which were included in Dodell-Feder and Tamir's (2018) meta-analysis. In principle, the results confirmed the evidential value. However, as statistical power was low, the evidential value was not robust.

Second, Mumper and Gerrig (2017) synthesized effects from 30 studies in which the frequency of reading fiction was correlated with social cognitive skills. Original empirical research was included if it represented a healthy adult sample (i.e., if participants were not recruited because they had been given a certain clinical diagnosis); if it examined a correlation between an indicator of lifetime print fiction and/or nonfiction exposure and a measure of social cognition (i.e., mentalizing or empathy); and if it was reported in English.

The volume of both fiction and nonfiction reading was found to be significantly positively correlated with empathy, aggregate *r*s = .07 and .05, respectively, and also with mentalizing, aggregate *r*s = .21 and .09. Moderator analyses examined whether these synthesized effects were impacted by the following variables: publication status and percentage of participating women as in Dodell-Feder and Tamir (2018), and additionally the type of experimental design – this variable reflected whether dependent measures were assessed only after a reading assignment or if no such assignment had taken place, given that this sort of task can affect dependent measures. None of these variables emerged as a significant moderator. Aside from the moderator analysis for publication status, Mumper and Gerrig (2017) did not examine the risk of publication bias.

Importantly, the synthesis by Dodell-Feder and Tamir (2018) pertains to effects of a single reading assignment typically comprising one short story or text excerpt. A single reading unit does not represent systematic training rooted in frequent engagement. Hence, the effects observed by Dodell-Feder and Tamir (2018) cannot be due to an acquisition of social-cognitive skills based on regular training (Eekhof et al., 2022; Mar, 2018a, 2018b). Other processes must be at work instead. For example, Mar has proposed that consuming a narrative, including reading a fictional story, could evoke a "social-processing mind-set" (Mar, 2018b, p. 259) or a "mentalizing mental mode" (Mar, 2018a, p. 469). These are states prompting better performance in subsequent tasks assessing social cognition. In the language of cognitive psychology, such facilitation of cognitive processes through prior activation is termed priming (see Lenhart & Richter, 2022). In the studies included in the meta-analysis by Mumper and Gerrig (2017), in contrast, participants were not usually tasked to read before their social-cognitive abilities were tested; and cases in which such assignments were involved were controlled for statistically (by including the type of experimental design as a moderator variable). Hence, the findings reported by Mumper and Gerrig (2017) do not

reflect processes of priming, but could in principle trace back to effects of repeated regular practice of social-cognitive skills through reading fiction.

Taken together, these two meta-analyses (Dodell-Feder & Tamir, 2018; Mumper & Gerrig, 2017) make important contributions to the literature as they show that reading fiction is accompanied by small-sized benefits for social cognition. However, although the body of research covered was adequate given the infancy of the field at the time, it is likely to have limited power. This may explain why neither meta-analysis detected any significant moderator variables. Furthermore, both meta-analyses were confined to the area of social cognition, so do not permit conclusions about broader cognitive effects and correlates of reading fiction.

The Present Meta-Analyses

We carried out two meta-analyses to significantly update and extend previous work. The hypotheses, methods, and analyses of both reviews were pre-registered at the Open Science Framework, https://osf.io/mskcz/?view_only=c6d5a6dbde2a4d0c88dbfdf5a72e30ab. Meta-Analysis 1 paralleled the review by Dodell-Feder and Tamir (2018), and Meta-Analysis 2 the one by Mumper and Gerrig (2017). We applied the same inclusion criteria as these authors with the following modifications: distinct from Mumper and Gerrig (2017) but in line with Dodell-Feder and Tamir (2018), we did not restrict samples to adults but included the full range of literate ages. Following Mumper and Gerrig (2017) but differing from Dodell-Feder and Tamir (2018), we did not include patient samples (note that our literature search did not detect any study that was excluded for working with a patient sample, so that ultimately this criterion did not have an effect; see Figure 1). Like Mumper and Gerrig (2017), we included studies reported in English only. Like Dodell-Feder & Tamir (2018), in Meta-Analysis 1 we included reading nothing and reading nonfiction as comparison conditions, in addition to any condition that did not involve reading fiction, such as watching fiction. This was deemed important to address the question whether potential effects are due to fiction

consumption, regardless of the activity of reading. Finally, in view of the controversies described (i.e., whether fiction-based effects on empathy and ToM are robust and/or linked with adverse effects; whether knowledge acquisition is a specific benefit of reading fiction or applies to reading nonfiction to a comparable or even higher degree; whether beliefs learned from fiction are on balance more likely to be correct than incorrect; and whether reading fiction is more likely linked with an increase or a reduction in biases and prejudices) we aimed to synthesize effects on all potentially longer-lasting cognitive outcomes reflecting either desirable or adverse traits rather than states. Following Cattell and Scheier (1961), we understand traits (e.g., creativity, verbal abilities) as temporally sustained behavioral dispositions that generalize across situations, whereas states (e.g., story-world absorption, identification with story characters) are thought to fluctuate both across time and situational circumstances. This means that we expanded our focus beyond social cognition. State-level indicators of cognition were not considered because they do not inherently mirror a more or less persistent cognitive benefit or disadvantage[1]. Direct measures of literacy were also excluded since we were interested in outcomes going beyond reading ability.

[1] This is in conflict with our discussion of Dodell-Feder and Tamir's (2018) meta-analysis where we questioned the persistence of their findings. In its core, our argument concerned the brevity of the reading assignments, not the measures used to assess social cognition. Nevertheless, trait-level indicators, such as the ones employed, should not be sensitive to short interventions, which in turn raises concerns as to whether those indicators are in fact set at the level of traits. Whilst we are unable to resolve this issue within the context of this paper, we restrict the focus of the current meta-analyses to indicators that are typically believed to reflect longer-lasting cognitive benefits or disadvantages, even if these may be generally considered as state-level indicators in the future.

We examined the same moderator variables as Dodell-Feder and Tamir (2018) and Mumper and Gerrig (2017), that is, sample type, type of comparison group, type of assessment measure (as in Dodell-Feder and Tamir's (2018) synthesis, this concerned whether measures were based on behavioral performance or on self-report), type of outcome, participant age, percentage of female participants, publication status, and type of experimental design, except for the following amendments: Dodell-Feder and Tamir's (2018) variable sample type (students *vs.* Amazon's Mechanical Turk workers) may reflect not just differing incentives, but also differences in participant supervision – students often complete study materials face-to-face with a researcher in a laboratory, whereas Amazon's Mechanical Turk workers by definition participate through an online survey, typically without researchers monitoring them. To allow us to disentangle effects of incentives and supervision, we implemented two moderator variables: one for type of participant group (students *vs.* community), and one for study format (in-person *vs.* online). For Meta-Analysis 1, we added the length of reading assignments as a moderator variable. If reading fiction achieves purported effects on cognition because the cognitive skills relevant to real life are practiced during reading, then the size of effects should grow with increasing length of reading assignments. If other processes drive these effects (for instance, if fiction-reading assignments serve to temporarily prime cognitive skills that are assessed immediately after the assignment), no moderating influence would be expected. For Meta-Analysis 2, we implemented the fictionality of print exposure as an explicit moderator. Hence, we were able to ascertain whether correlations with cognitive outcomes are stronger for fiction than for nonfiction exposure. In contrast, Mumper and Gerrig (2017) ran separate analyses for fiction and nonfiction exposure, which did not allow such direct comparisons.

The following research questions were investigated in Meta-Analysis 1:

1.1 How many experimental studies have investigated cognitive effects of reading fiction?

1.2 What is the synthesized effect size across these experiments?

1.3 Does this total effect size depend on the moderator variables outlined above (i.e., type of assessment measure, publication status, study format, type of participant group, outcome variable, type of comparison group, participant age, percentage of female participants, length of reading assignment)?

Meta-Analysis 2 was conducted to answer the following research questions:

2.1 How many correlational studies have investigated cognitive effects of reading fiction?

2.2 What is the synthesized effect size across these studies?

2.3 Does this total effect size depend on the moderator variables outlined above (i.e., type of assessment measure, publication status, study format, type of participant group, outcome variable, fictionality of print exposure, study design, participant age, percentage of female participants)?

Materials and Method

Due to a high degree of overlap regarding materials and method, this section is reported jointly for the two meta-analyses.

Transparency and Openness

All meta-analytic data, analysis code, and research materials (including our coding scheme) are available at https://osf.io/dus2w/?view_only=bb8790263de94ccfa24eecf3be5c3833. This project was pre-registered at the Open Science Framework, https://osf.io/mskcz/?view_only=c6d5a6dbde2a4d0c88dbfdf5a72e30ab.

Selection Criteria

Meta-Analysis 1

We considered research that meets the following inclusion criteria:

- it follows a true experimental design with random assignment to condition;
- it includes a control condition that does not involve reading fiction;

- it includes at least one measure of cognition that goes beyond both literacy and state-level indicators; following the APA dictionary of psychology (VandenBos, 2015), we defined cognition as all forms of knowing and awareness, such as perceiving, conceiving, remembering, reasoning, judging, imagining, and problem-solving;
- full text is written in English.

The only exclusion criterion was:

- research that focuses on patient samples (however, we included data from non-clinical or control groups where appropriate).

Meta-Analysis 2

We included research that fulfills the following criteria:

- it correlates at least one measure of lifetime exposure to print fiction or nonfiction with at least one measure of cognition that goes beyond both literacy and state-level indicators; cognition was defined as in Meta-Analysis 1;
- full text is written in English.

We applied the following exclusion criterion:

- research that focuses on patient samples.

Reports that did not contain sufficient information to assess eligibility were excluded. If reports did not include sufficient statistical information for analysis or if this information could not be calculated from publicly shared data, authors with findable current contact details whose reports were published within the past ten years were contacted. In cases where the necessary details still could not be obtained, reports were excluded.

Search Strategy

We performed a joint literature search for both meta-analyses. On the 21st of April 2022, a database search was carried out using PsycINFO (via EBSCOhost) and Web of Science (via Clarivate), in each case with the search term "fict* OR non-fict* OR nonfict*" (no specific search fields selected). To retrieve unpublished work, the same search term was

used on the PsyArXiv Preprints server (https://psyarxiv.com; 30th of August 2022) and on the following dissertation repositories: Dissertations Express (https://dissexpress.proquest.com/search.html; 30th of August 2022), Open Access Theses and Dissertations (https://oatd.org; 31st of August 2022), and Theses Canada Portal (https://library-archives.canada.ca/eng; 30th of August 2022). Our pre-registration additionally included the repository OpenThesis (http://www.openthesis.org); however, this website was not accessible at the time of data collection, so was not included as an information source. We also searched the publication lists of authors whose studies were included (at least one author of each report was considered). In addition, we contacted these authors for unpublished studies, and put out a call for unpublished studies via the mailing lists of the Society for Text and Discourse (ST&D; 7th of September 2022) and the International Society for the Empirical Study of Literature (IGEL; 14th of September 2022). Furthermore, we searched the reference lists of articles meeting inclusion criteria and also the reference lists of the reviews/meta-analyses by Appel et al. (2021), Barnes (2018), Best (2020, 2021), Black et al. (2021), Burke et al. (2016), Dodell-Feder and Tamir (2018), Eekhof et al. (2022), Goffin and Friend (2022), Hopkins and Weisberg (2017), Jacobs and Willems (2018), Koopman (2018), Mar (2018b), Mumper and Gerrig (2017), Oatley et al. (2018), and Wimmer and Ferguson (2022). We also considered research reports found through manual search. Three authors were contacted since full-text reports did not include all statistics needed to calculate effect sizes; two replied, with one providing the requested details.

Data Extraction and Coding of Moderator Variables

All steps reported within this section were carried out by the first author. Titles and abstracts of the reports retrieved through database search (i.e., involving PsycINFO and Web of Science) were screened using Abstrackr (Wallace et al., 2012). This tool utilizes machine-learning algorithms trained by researchers' screening decisions in order to present citation abstracts in such a sequence that those with the highest probability of meeting inclusion

criteria appear first within the screening process. Tsou et al. (2020) demonstrated that Abstrackr is effective for this purpose. They further suggested that when adopting a conservative approach, it is feasible to stop screening after 70 to 80 percent of reports have been examined. In line with this suggestion, screening was terminated after 22,661 out of 31,533 possible inspections, corresponding to a screening rate of 71.86 percent. Records identified through other sources were screened based on titles and additionally on abstracts if the title was compatible with the topic of the current meta-analyses. To further assess eligibility of the entries remaining after screening, each full text was examined. The final set of reports used for analyses was approved by all authors.

The following information was extracted from reports for both meta-analyses:

- author names
- year of publication
- title
- publication status (published *vs.* unpublished) – research was regarded as published if respective reports were released by a publishing company, for example as journal articles or chapters in edited books, and otherwise as unpublished, for example dissertations and manuscripts submitted to but not accepted for publication in a journal.
- study format (online *vs.* in person) – this variable was primarily meant to reflect the degree of participant supervision, which is typically relatively high in research carried out in person and relatively low in online research. Consequently, study formats within which participants completed paper-pencil-based measures remotely, that is, without researcher supervision, were coded as online.
- types of participant groups (students *vs.* community) – this variable was geared towards the distinction between convenience samples, which in psychological research routinely consist of students who receive course credits in exchange for participation, and non-convenience samples, who are recruited from the general public and often receive

monetary compensation. In line with this, school pupils were coded as community samples, as were samples that included students as a minority.

- sample size – split by experimental group for Meta-Analysis 1 and total sample size for Meta-Analysis 2.
- participant age in years (mean and standard deviation) – split by experimental group for Meta-Analysis 1 and for the total sample for Meta-Analysis 2.
- percentage of female participants – split by experimental group for Meta-Analysis 1 and for the total sample for Meta-Analysis 2.
- outcome variable – groupings served three aims: first, that subcategories were in line with contemporary cognitive concepts; second, that the subcategories were distinct; third, that the number of outcome variables that could be assigned to one of the subcategories was maximized. This was meant to maximally exhaust the current pool of studies in the moderator analyses to test for effects of different outcome variables. As a result, the following subcategories were used for Meta-Analysis 1: empathy (a collective term for processes conducive to sharing and comprehending others' affects; Håkansson Eklund & Summer Meranius, 2021); knowledge (understood as true belief, as in Allen et al., 1992, for example); mentalizing (understanding mental states; Preston et al., 2020); moral cognition (thought processes relevant to morality, i.e., the repertoire of customs and values prevalent in a given society to govern social conduct; Moll et al., 2005); outgroup judgments (judgments about people from a group that participants do not belong to or identify with; VandenBos, 2015); and other thinking processes (cognitive processes not covered by the aforementioned subcategories). Meta-Analysis 2 was based on the following subcategories: creativity (the capacity to generate novel and useful products; Mumford, 2003); general cognitive abilities (processes contributing to reasoning, planning, problem solving, abstract thinking, understanding complex ideas, learning fast and from experience; Gottfredson, 1997); verbal abilities (the capacity to apprehend and

communicate adequately with words; VandenBos, 2015); and empathy, mentalizing, moral cognition, outgroup judgments, and other thinking processes (as defined as in Meta-Analysis 1).

- name of assessment measure (e.g., Interpersonal Reactivity Index [IRI; Davis, 1980]).
- results – for Meta-Analysis 1, we extracted means and standard deviations of the dependent measure, split by experimental group, or alternative statistics if these were not reported; for Meta-Analysis 2, we used bivariate correlations between lifetime exposure to print fiction or nonfiction with each cognitive outcome.
- type of assessment measure which specifically targeted the use of self-report measurements. These assessments are prone to biases such as socially desirable responding, a limitation meant to be circumvented in other measures, especially those based on behavioral performance (e.g., Moosbrugger & Kelava, 2020). Hence, risk of bias was deemed to be higher for self-report than for performance-based measures. For Meta-Analysis 1, we evaluated each dependent measure according to whether it relied on self-report or behavioral performance. For Meta-Analysis 2, the assessment of both the independent variable (i.e., lifetime print exposure) and the dependent variable(s) was taken into account. The levels of this moderator variable were, first, both the outcome and print exposure were assessed via self-report, second, either the outcome or print exposure was assessed via self-report, and third, neither the outcome nor print exposure was assessed via self-report.

For Meta-Analysis 1, we additionally extracted the following information:

- type of condition that was compared to reading fiction – the current pool of studies contained reading nonfiction, reading nothing, and watching fiction.
- length of the reading assignment as word count; if the precise word count was not reported, it was estimated based on numbers provided at https://self-publishingschool.com/how-many-words-in-a-novel/ .

The following information was extracted for Meta-Analysis 2 only[2]:

- Study design (correlational *vs.* experimental) – experimental designs referred to studies in which cognition was assessed after a reading assignment, whereas no such assignments were set within correlational designs.

The coding scheme is available at https://osf.io/dus2w/?view_only=bb8790263de94ccfa24eecf3be5c3833.

Statistical Analysis

Effect size calculations were based on the formulas provided in Borenstein (2009) and Lipsey and Wilson (2001) as well as the *esc* package in R (R version 4.2.3). All further meta-analytical steps were carried out using the R *metafor* package (Viechtbauer, 2010). If not otherwise mentioned, we adopted the standard 5% significance level for inferential tests. Meta-analytical code can be accessed at https://osf.io/dus2w/?view_only=bb8790263de94ccfa24eecf3be5c3833. For Meta-Analysis 1, effect sizes were calculated as bias-corrected Hedges' *g*, representing the standardized mean difference between reading fiction and the comparison group, such that positive effect sizes represent better performance in the fiction group. In order to replicate Dodell-Feder and Tamir's (2018) approach, we used raw unadjusted means and standard deviations, and compared only post-reading between-groups scores. For Meta-Analysis 2, effect sizes were obtained by transforming the correlations from each contributing study into a Fisher's *Z* score, then averaging the *Z* scores together, and transforming the averaged *Z* score back into an aggregate correlation for ease of interpretation. Positive correlations indicate a positive linear relationship between lifetime exposure to written fiction and cognitive benefits. If cognitive

[2] Our pre-registration included fiction genre as a moderator variable. However, this could not be put into effect since the vast majority of studies reported general fiction scores only but did not differentiate particular fiction genres.

skills were assessed more than once within a study, we utilized only the correlation between print exposure and the first cognitive assessment.

Effect sizes were weighted by their inverse variance within both meta-analyses. We observed two types of dependencies in our data: first, that more than one dependent measure was assessed in the same sample, and second, that the same group of participants was compared more than once with another group of participants on the same dependent measure. Thus, we implemented multilevel random-effects meta-analytic models that account for variance at the different levels, in particular, variance in the observed effect sizes (level 1), variance between effect sizes within a study (level 2), and variance between studies (level 3). Since there were cross-level dependencies, we generated cluster-robust standard errors, statistical tests, and confidence intervals on estimates from the multi-level meta-analytic models. To gauge heterogeneity, we calculated χ^2 (Q) and Higgins I^2 (Harrer et al., 2022) statistics.

The robustness of our findings was assessed by the following sensitivity analyses (paralleling the approach taken by Dodell-Feder & Tamir, 2018): first, we examined effect sizes for influential outliers, defined as effect sizes with standardized residual values exceeding 3.0 (Cohen & Cohen, 2003) whose Cook's Distance values exceeded $4/(n\text{-}k\text{-}1)$ (Fox, 1991). Second, we employed a leave-one-out procedure to gauge the impact of each individual effect size (i.e., rerunning the multilevel model leaving out one effect size at a time) and study (i.e., rerunning the multilevel model leaving one study out at a time) on the overall effect and amount of heterogeneity.

The impact of moderator variables was tested in a series of meta-regression models, with each model including one moderator variable. Finally, the following tests ascertained the risk of publication bias: we included publication status (published *vs.* unpublished) as a moderator within meta-regression models; we also compiled funnel plots and conducted

respective Egger's tests. In addition to these pre-registered analyses, we carried out *p*-curve analyses using the *p*-curve app (version 4.6, http://www.p-curve.com/app4/).

Results

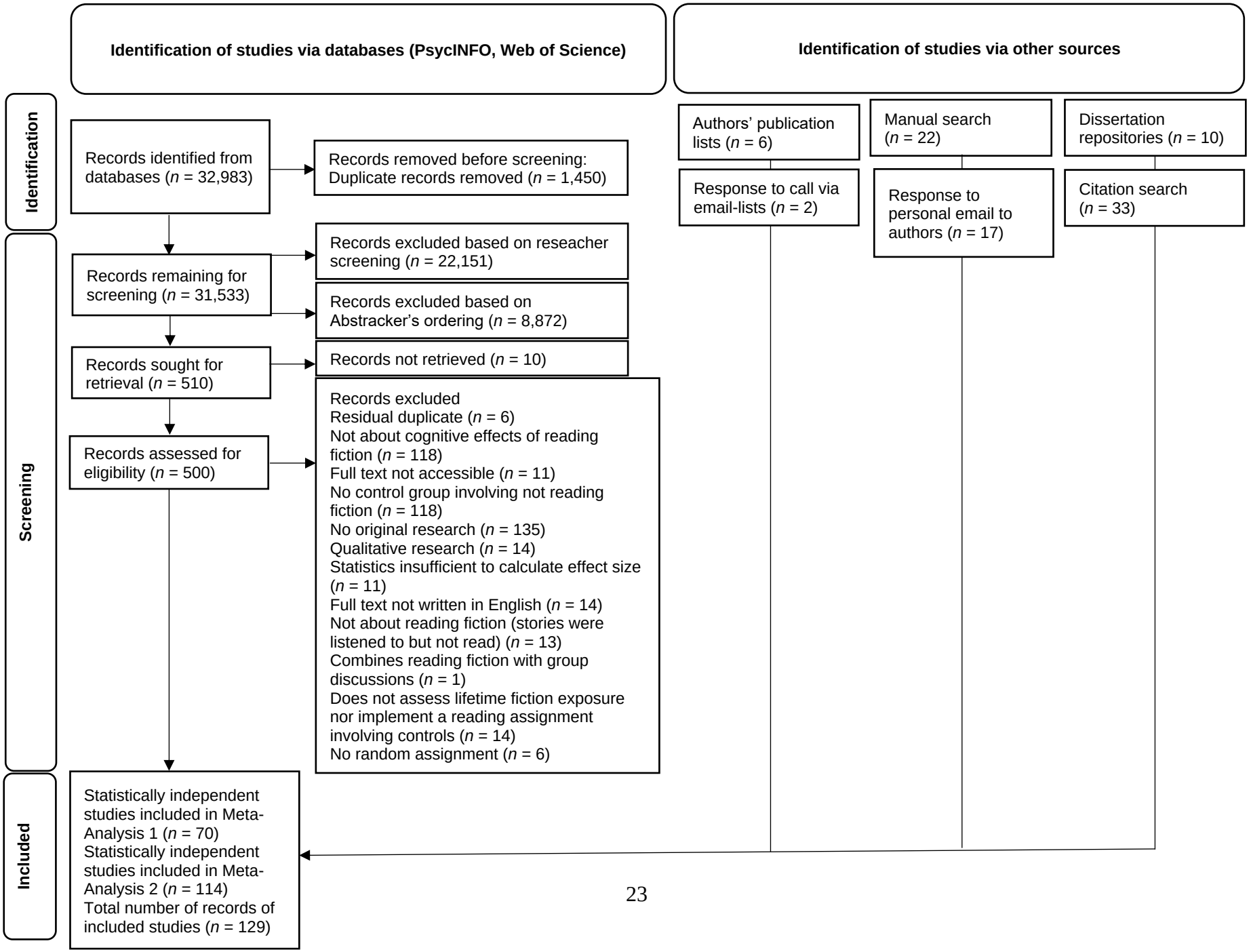
To follow the conventions of multi-study articles, the statistical results of Meta-Analysis 1 are reported before the results of Meta-Analysis 2.

Meta-Analyses 1 and 2: Number of Studies

Our database search detected 32,983 entries on the 21st of April 2022. During the initial selection process, duplicates were removed. After titles and abstracts were screened, full texts of the remaining records were sought and checked for eligibility. Figure 1 provides detailed information on the number of entries excluded at each step. Finally, 39 records were found eligible for one or both meta-analyses. A further 90 reports were detected through means other than database search (as outlined above, namely authors' publication lists, manual search, dissertation repositories, responses to calls via email lists, authors' responses to personal emails, and search of reference lists). In sum, 129 reports were included in the present meta-analyses, 12 of which were included in both meta-analyses. Meta-Analysis 1 included 70 and Meta-Analysis 2 included 114 statistically independent studies.

Figure 1

PRISMA Flowchart of the Literature Search


Identification of studies via databases (PsycINFO, Web of Science)
Identification of studies via other sources
Identification
Screening
Included
Records identified from databases (n = 32,983)
Records removed before screening: Duplicate records removed (n = 1,450)
Records remaining for screening (n = 31,533)
Records excluded based on reseacher screening (n = 22,151)
Records excluded based on Abstracker's ordering (n = 8,872)
Records sought for retrieval (n = 510)
Records not retrieved (n = 10)
Records assessed for eligibility (n = 500)
Records excluded
Residual duplicate (n = 6)
Not about cognitive effects of reading fiction (n = 118)
Full text not accessible (n = 11)
No control group involving not reading fiction (n = 118)
No original research (n = 135)
Qualitative research (n = 14)
Statistics insufficient to calculate effect size (n = 11)
Full text not written in English (n = 14)
Not about reading fiction (stories were listened to but not read) (n = 13)
Combines reading fiction with group discussions (n = 1)
Does not assess lifetime fiction exposure nor implement a reading assignment involving controls (n = 14)
No random assignment (n = 6)
Authors' publication lists (n = 6)
Manual search (n = 22)
Dissertation repositories (n = 10)
Response to call via email-lists (n = 2)
Response to personal email to authors (n = 17)
Citation search (n = 33)
Statistically independent studies included in Meta-Analysis 1 (n = 70)
Statistically independent studies included in Meta-Analysis 2 (n = 114)
Total number of records of included studies (n = 129)

Meta-Analysis 1

Characteristics of Studies and Samples

The current pool of studies involved 5,640 participants assigned to read fiction and 5,532 participants allocated to a control condition. For individual studies, sample sizes of groups assigned to read fiction varied between n = 12 (Hakemulder, 2000, Study 2) and n = 342 (Panero et al., 2016; Kidd & Castano, 2017a), inter-quartile range (*IQR*): [30.25, 78.75], whilst those of control groups varied between n = 12 (Hakemulder, 2000, Study 2) and n = 189 (Panero et al., 2016; Kidd & Castano, 2017a), *IQR*: [30, 75.5].

Mean ages ranged from 5.55 (Justice et al., 2005) to 44.00 years (Maxim, 2022, sample 2), *IQR*: [19.80, 34.20]. Regarding participant gender, percentage of female respondents varied between 38.20% (Vezzali et al., 2021, Experiment 2) and 87.00% (Lenhart & Richter, 2022, Experiment 2), *IQR*: [52.60%, 76.37%]. 71.43% of studies were published (*vs.* unpublished), 55.88% were conducted in-person (*vs.* online), and 61.43% worked with a community (*vs.* student) sample. For 76.20% of effect sizes, reading fiction was compared to reading nonfiction, whilst 19.83% of effect sizes concerned comparisons with reading nothing, and 3.97% of effect sizes relied on comparisons with watching fiction. The length of the reading assignment ranged from 273 (Kilaru et al., 2014) to 350,100 words (Vezzali et al., 2021, Experiment 2), *IQR*: [1,500, 6,000]. In terms of the outcome variables, empathy was addressed by 28.65% of comparisons, knowledge by 11.89%, mentalizing by 21.62%, moral cognition by 21.35%, outgroup judgments by 14.05%, and other thinking processes by 2.43% (for exemplary measures of each outcome variable, see Table 1). Dependent measures were based on behavioral performance for 56% of comparisons and on self-report for the remaining 44%.

Table 1

Selection of Cognitive Measures

Outcome variable	Measure	Abbreviation	Description	Method of assessment	Meta-Analysis 1: original studies using measure	Meta-Analysis 2: original studies using measure
Creativity	Versions of the Alternate Uses Task (Dippo, 2013; George & Wiley, 2019; Guilford, 1967)	AUT	List creative uses of common objects	Behavioral performance	N/A	Black & Barnes (2021); Wimmer et al. (2022b)
	Behavioral Identification Form (Vallacher & Wegner, 1989)	BIF	Choose one of two alternative descriptions (one lower and one higher in level) of an act	Behavioral performance	N/A	Stanisic (2019)
	Insight problems taken from Förster et al. (2004)	N/A	Solve problems likely to provoke an impasse the conquest of which is typically described as "aha" experience	Behavioral performance	N/A	Stanisic (2019)
Empathy	Reading the Mind in the Eyes Test - Revised (Baron-Cohen et al., 2001)	RMET	Selecting one of several response options, determine actors' mental state from photograph presenting only their eye region	Behavioral performance	Black & Barnes (2015); DeMulder et al. (2017); Dodell-Feder et al. (2022); Kidd & Castano (2013); Lenhart & Richter (2022); and others	Castano et al. (2020); Kidd & Castano (2017); Samur et al. (2017); Schwerin & Lenhart (2022); Takahashi et al. (2023)
	Interpersonal Reactivity Index (Davis, 1980)	IRI	Using a rating scale, indicate the following tendencies: take others' perspective, feel empathic concern for others, transpose oneself into fictitious characters, feel distressed in tense interpersonal settings	Self-report	Johnson et al. (2013); Pigden (2021); Savitri & Nuha (2021); Seddio (2017);	Eekhof et al. (2021); Liu & Want (2015); Tabullo et al. (2018); Waitz et al.

Outcome variable	Measure	Abbreviation	Description	Method of assessment	Meta-Analysis 1: original studies using measure	Meta-Analysis 2: original studies using measure
					Walkington et al. (2020); and others	(2015); Whalen (2010)
	Eye-tracking paradigm on sad vs. emotionally neutral videos	N/A	Watch sad vs. emotionally neutral videos while participants' pupil size and fixations on the actress' eye-region are tracked	Implicit measure/not self-report	Wimmer et al. (2022a)	Wimmer et al. (2022a)
General cognitive abilities	Versions of verbal fluency tasks (see, e.g., Sincoff & Sternberg, 1987)	N/A	Produce as many words of a given category as possible within a fixed time frame	Behavioral performance	N/A	Cunningham & Stanovich (1991); Stanovich & Cunningham (1992); Tabullo & Pithod (2020)
	Grade Point Average	GPA	Average of academic grades achieved	Behavioral performance	N/A	Spear-Swerling et al. (2020); Tabullo & Pithod (2020)
	Versions of matrices tests (e.g., Raven, 1962)	N/A	Indicate the option that completes a pattern arranged in a matrix	Behavioral performance	N/A	Brysbaert et al. (2020); Grant (2011); Lenhart et al. (2023); Stanovich & Cunningham (1992)
Knowledge	Science knowledge about dinosaurs	N/A	Select the single correct response regarding dinosaurs	Behavioral performance	Willis (1998)	N/A
	Autism Stigma and Knowledge Questionnaire (Harrison et al., 2017)	ASK-Q	State whether one agrees with, disagrees with, or is uncertain about the accuracy of statements on ASD in terms of diagnosis, etiology, treatment, and stigma	Behavioral performance	Stern (2020)	N/A

Outcome variable	Measure	Abbreviation	Description	Method of assessment	Meta-Analysis 1: original studies using measure	Meta-Analysis 2: original studies using measure
	Recall of Opioid prescription guidelines	N/A	Write down all information one can recall from a reading without reviewing it	Behavioral performance	Kilaru et al. (2014)	N/A
Mentalizing	Yoni task (Shamay-Tsoory & Aharon-Peretz, 2007)	N/A	Identify the item Yoni (cartoon outline of a face) refers to based on a sentence and specific cues (e.g., Yoni's eye gaze and/or facial expression)	Behavioral performance	DeMulder et al. (2017); Kidd & Castano (2013); Turner & Valée-Tourangeau (2020)	DeMulder et al. (2017); Turner & Valée-Tourangeau (2020); van Kuijk et al. (2018)
	Interpersonal Perception Task-15 (Costanzo & Archer, 1993)	IPT-15	Having watched a video vignette involving social interactions, select the correct statement about the interaction based on inferences from characters' dynamic non-verbal cues	Behavioral performance	None	Bischoff (2013); Mar et al. (2006)
	Revised version of the Frith-Happé animations task (White et al., 2011)	N/A	Responding to multiple-choice questions, ascribe mental states/interactions to animated triangles	Behavioral performance	Wimmer et al. (2022a)	Wimmer et al. (2021b)
Moral cognition	Moral Identity Questionnaire (Black & Reynolds, 2016)	N/A	On a Likert scale, rate the importance morality plays in one's identity and the emphasis one puts on acting in accordance with moral principles	Self-report	None	Black & Barnes (2021c)
	Versions of the Implicit Association Test (Greenwald et al., 1998)	IAT	Respond to a concept with the same key used to respond to either a moral or immoral concept; the more the key pairings with immoral concepts require longer response latencies than key pairings with moral concepts, the higher one's level of moral cognition	Implicit measure/behavioral performance	Pigden (2021); Wimmer et al. (2022a)	Redman (2017); Wimmer et al., (2021a, b)

Outcome variable	Measure	Abbreviation	Description	Method of assessment	Meta-Analysis 1: original studies using measure	Meta-Analysis 2: original studies using measure
	Version of the Moral Judgment Task (e.g., Young et al., 2012)	MJT	Using a rating scale, assess the moral permissibility of actions varying in the valence of their actual results (neutral vs. negative) and in the valence of their results as expected by the agent (neutral vs. negative)	Behavioral performance	Dodell-Feder et al. (2022); Kidd & Castano (2019)	Black & Barnes (2021); Xu et al. (2022)
Outgroup judgments	Ambivalent Sexism Inventory (Glick & Fiske, 1997)	ASI	Using a rating scale, indicate the degree to which one agrees with statements reflecting hostile and benevolent sexism	Self-report	Redman (2017)	Redman (2017)
	Word fragment completion task	N/A	Complete word fragments that have either a positive and a neutral completion or a negative and a neutral completion. Positive outgroup judgments are reflected by a high number of positive vs. neutral completions; negative outgroup judgments are reflected by a high number of negative vs. neutral completions	Implicit measure/behavioral performance	Johnson et al. (2013)	None
	Genderism and Transnegativity Scale (Tebbe et al., 2014)	GTS-R	Using a rating scale, state the extent to which one agrees with statements reflecting negative attitudes and the inclination towards violence towards transgender individuals	Self-report	Orellana et al. (2022)	None
Verbal abilities	Sentence completion task (items taken from the reading section of the	N/A	Choosing from multiple response options, decide how to best complete a sentence	Behavioral performance	N/A	Mar & Rain (2015); Martin-Chang et al. (2021)

Outcome variable	Measure	Abbreviation	Description	Method of assessment	Meta-Analysis 1: original studies using measure	Meta-Analysis 2: original studies using measure
	Scholastic Aptitude Test [SAT])					
	Yes/No Vocabulary Test Dutch (Lemhöfer & Broersma, 2012)	N/A	Indicate whether or not a string of letters is a word in a given language	Behavioral performance	N/A	Brysbaert et al. (2020)
	Vocabulary subtest of the Nelson-Denny Reading Test (Brown et al., 1993)	N/A	State the meanings of words selecting one of multiple response options each	Behavioral performance	N/A	Grant (2011); Stanovich & Cunningham (1992); Vermeiren et al. (2022)
Other thinking processes	Adapted version of the Memories Experiences Questionnaire (Bakhtiari et al., 2022; Østby et al., 2012)	MEQ	Using a rating scale, rate the vividness of a retrieved memory	Self-report	Maxim (2022)	None
	Adapted version of the Relationships Beliefs Inventory (Eidelson & Epstein, 1982)	RBI	Using a rating scale, rate the degree to which one supports maladaptive beliefs about romantic relationships, namely that disagreement is destructive, that mindreading is expected, that partners cannot change, that the sexes are different, and that sexual perfectionism is expected	Self-report	None	Stern et al. (2019)

Outcome variable	Measure	Abbreviation	Description	Method of assessment	Meta-Analysis 1: original studies using measure	Meta-Analysis 2: original studies using measure
	Essentialism Scale (Bastian & Haslam, 2006)	N/A	Using a rating scale, indicate the extent to which one believes that human attributes are immutable in the sense that individuals fall into discrete categories, that differences between individuals permit many inferences regarding further differences, and that human attributes have a biological basis	Self-report	None	Castano et al. (2021)

Note. Three exemplary measures per outcome variable, meant to show the diversity of operationalizations across studies, are provided. The RMET is categorized as a measure of empathy, whereas in most other research, including Dodell-Feder and Tamir's (2018) meta-analysis, it is treated as a measure of mentalizing. The current classification is based on evidence suggesting that the RMET indicates emotion recognition rather than mentalizing (Oakley et al., 2016), and on the common classification of emotion recognition as a component of empathy (see Introduction).

Meta-Analysis and Robustness Checks

Our primary research question addressed whether and to what extent reading fiction, compared with control activities, enhances cognition. According to our meta-analysis of 371 effect sizes from 70 studies, compared with control activities, reading fiction led to greater cognitive benefits. This effect was small in size but statistically significant, g = 0.14, 95% *CI* [0.06, 0.21], p = .0004. The presence of heterogeneity was indicated by a significant Q statistic, $Q(370) = 1086.50$, $p < .0001$. Furthermore, total I^2 was 72.11%, indicating a substantial amount of true variance (*vs.* sampling error) in effect size estimates. The majority of this came from the between-study cluster, I^2_{Level3} = 52.38%, compared with the within-study cluster, I^2_{Level2} = 19.72%. In sum, these findings suggest that reading fiction leads to small-sized cognitive benefits, and that the effect sizes differ systematically between studies due to factors that vary between studies (e.g., participant age, type of assessment measure). Robustness of the effect was confirmed using sensitivity analyses: standardized residual values ranged from −1.22 to 1.62, hence all figures were below the critical cut-off of 3.0 (as absolute value; Cohen & Cohen, 2003). Using the leave-one-out procedure at the effect size level, aggregate effect size estimates varied between g = 0.13 and g = 0.14, whilst total I^2 was between 71.33% and 72.77%. For the leave-one-out procedure at the study level, aggregate effect size estimates ranged from g = 0.12 to g = 0.15, and total I^2 from 70.22% to 73.79%.

Moderator Analyses

A series of moderator analyses tested whether the significant amount of heterogeneity could be explained by two types of factors: categorical (e.g., type of assessment measure) and continuous variables (e.g., sample age). A summary of the overall meta-analytic effect as well as the effects of these moderator variables is given in Table 2. For categorical moderators with up to three levels, we ran meta-regression models with an intercept. In this case, the p-value of the corresponding effect size parameter (see Table 2) indicates whether the effect of

the intercept differs significantly from zero. The effect of each of the subsequent moderator levels is compared to the intercept. Hence, the *p*-values of these effects reflect whether the effect of the moderator level differs significantly from that of the intercept, and so suggest whether the respective variable exerts a moderating influence. For categorical moderators with more than three levels (e.g., outcome variable), we ran meta-regression models without an intercept. Here, the effect of each moderator level is tested against zero, as reflected in the corresponding *p*-values. This was followed up with pairwise contrasts (reported in the text, not in Table 2), for which the significance level was adjusted using the Bonferroni method. For continuous moderators, the effect size parameter is a regression weight *b*, so that *p*-values indicate whether there is moderation in terms of a significant linear relationship between moderator and outcome.

According to the meta-regressions, the overall meta-analytic effect was moderated significantly by, first, type of comparison group. Specifically, effect sizes were greater when reading fiction was compared with either watching fiction, $g = 0.31$, or reading nothing, $g = 0.23$, than when reading fiction was compared with reading nonfiction, $g = 0.10$ (ps $< .006$). A second moderating effect emerged for outcome variable: significant effects were observed only when the outcome was either empathy, $g = 0.16$, or mentalizing, $g = 0.15$, whereas effects did not differ significantly from zero for the remaining outcomes (i.e., knowledge, moral cognition, outgroup judgments, and other thinking processes; ps $> .11$). Pairwise comparisons failed to reveal significant differences between effects of individual outcome variables (ps $> .07$, with $p_{crit} = .003$). There were no further significant moderations (ps $\geq .05$).

Publication Bias

The following analyses of publication bias included published work only. Visual inspection of the funnel plot revealed some degree of asymmetry as there were more data points to the right than to the left of the mean effect (see Figure 2A). Yet, although the slope of Egger's regression test for funnel plot asymmetry was positive, $b = 2.25$, $SE = 1.76$, $t(48) =$

1.28, $p = .21$, it was not significant. That is, the precision of the measured effect was not significantly linked with the magnitude of the effect.

According to *p*-curve analysis (see Figure 3A), the right skewness of the *p*-curve was significant, $Z = -10.99$, $p < .001$ (full *p*-curve) and $Z = -9.59$, $p < .001$ (half *p*-curve). Therefore, the current pool of studies can be thought to have evidential value that reading fiction leads to cognitive benefits. Additionally, the *p*-curve did not suggest evidential inadequacy (i.e., flatter than 33% power), $Z = 4.98$, $p > .999$ (full *p*-curve) and $Z = 11.72$, $p > .999$ (half *p*-curve). The estimated power of the tests included in the *p*-curve was 71%.

Meta-Analysis 2

Characteristics of Studies and Samples

The current pool of studies involved 30,503 participants. Sample sizes of individual studies varied between $n = 33$ (Stansfield & Bunce, 2014) and $n = 4{,}775$ (Lenhart et al., 2023), *IQR*: [104.00, 262.80]. Mean ages ranged from 8.67 (Arnold et al., 2019) to 59.29 years (Wimmer & Ferguson, 2022; Wimmer et al., 2021a), *IQR*: [19.64, 31.98]. Regarding participant gender, percentage of female volunteers varied between 39.68% (Allen et al., 1992) and 100% (Diekman et al., 2000; Study 1), *IQR*: [55.62%, 76.85%]. 74.56% of studies were published (*vs.* unpublished), 45.63% were conducted in-person (*vs.* online), and 50.47% worked with a community (*vs.* student) sample. The study design of 86.92% of comparisons was correlational (*vs.* experimental). Turning to outcome variables, creativity was targeted by 2.16% of comparisons, empathy by 43.17%, general cognitive abilities by 12.05%, mentalizing by 6.29%, moral cognition by 13.31%, outgroup judgments by 4.50%, verbal abilities by 8.99%, and other thinking by processes by 9.53% (for exemplary measures of each outcome variable see Table 1). Neither the outcome nor print exposure was assessed via self-report for 44.17%, either the outcome or print exposure was assessed via self-report for 49.37%, and both the outcome and print exposure were assessed via self-report for the remaining 6.46% of comparisons.

Meta-Analysis and Robustness Checks

The primary research question concerned whether and to what extent lifetime exposure to written fiction is associated with cognitive benefits. Our meta-analysis of 559 effect sizes from 114 studies revealed that there is a significant small-sized, positive association between lifetime exposure to written fiction and cognitive abilities, r = .16, 95% *CI* [.13, .19], $p < .0001$. Presence of heterogeneity was indicated by a significant Q statistic, $Q(558) = 5825.12$, $p < .0001$. In addition, total I^2 was 91.33%, indicating a substantial amount of true variance (*vs.* sampling error) in effect size estimates. The majority of this was rooted in the within-study cluster, I^2_{Level2} = 54.56%, compared to the between-study cluster, I^2_{Level3} = 36.78%. In sum, these findings suggest that lifetime exposure to written fiction is associated with small-sized cognitive benefits, and that the effect sizes differ systematically between studies due to factors that vary within studies (e.g., outcome variable). Robustness of the effect was confirmed using sensitivity analyses: standardized residual values ranged from -0.80 to 1.93, thus all figures were below the critical cut-off of 3.0 (understood as absolute value; Cohen & Cohen, 2003). Using the leave-one-out procedure at the effect size level, effect size estimates varied between r = .16 and r = .17, whilst total I^2 was between 90.10% and 91.53%. For the leave-one-out procedure at the study level, effect size estimates ranged from r = .16 to r = .17, and total I^2 from 90.08% to 91.68%.

Table 2

Meta-Analytic Results

		Meta-Analysis 1								Meta-Analysis 2							
				Effect Size Estimate								Effect Size Estimate					
		Number of Studies	**Number of ES**	**Parameter**	**95% *CI***	***SE***	***t***	***p***	***p* informs whether …**	***Q***	**Number of Studies**	**Number of ES**	**Parameter**	**95% *CI***	***SE***	***t***	***p***	***p* informs whether …**	***Q***
Overall Estimate		70	371	*g* = 0.14	[0.06, 0.21]	0.04	3.75	.0004	… *g* is different from 0	1086.50***	114	559	*r* = .16	[.13, .19]	0.02	10.33	< .0001	… *r* is different from 0	5825.12***
Categorical Moderators																			
Type of assessment measure	Behavioral performance (Meta-analysis 1) / neither print exposure nor outcome assessed via self-report (Meta-Analysis 2) (intercept)		206	*g* = 0.11	[0.03, 0.19]	0.04	2.75	.0076	… *g* is different from 0	1041.26***		246	*r* = .19	[.15, .24]	0.02	-8.33	<.0001	… *r* is different from 0	5723.39***
	Either print exposure or outcome assessed via self-report (*vs.* intercept)											275	*r*= .13	[.08, .18]	0.03	-2.28	.0245	… effect of moderator is different from 0	
	Self-report (Meta-analysis 1) / Both print exposure and outcome assessed via self-report		165	*g* = 0.18	[0.09, 0.28]	0.05	1.55	.1265	… effect of moderator is different from 0			36	*r* = .17	[.08, .27]	0.05	-0.35	.7289	… effect of moderator is different from 0	

		Meta-Analysis 1									Meta-Analysis 2								
				Effect Size Estimate					*p* informs whether …	*Q*			Effect Size Estimate					*p* informs whether …	*Q*
		Number of Studies	**Number of ES**	**Parameter**	**95% *CI***	***SE***	***t***	***p***			**Number of Studies**	**Number of ES**	**Parameter**	**95% *CI***	***SE***	***t***	***p***		
	(Meta-analysis 2) (*vs.* intercept)																		
Publication status	Published (intercept)	50	267	g = 0.18	[0.09, 0.27]	0.05	3.91	.0002	… *g* is different from 0	1036.65***	85	428	r = .18	[.14, .22]	0.02	9.44	< .0001	… *r* is different from 0	5679.33***
	Unpublished (*vs.* published)	20	104	g = 0.04	[−0.11, 0.18]	0.07	-1.98	.0512	… effect of moderator is different from 0		29	131	r = .11	[.05, .17]	0.03	-2.04	.0436	… effect of moderator is different from 0	
Study format	Online (intercept)	30	127	g = 0.07	[−0.00, 0.15]	0.04	1.89	.0630	… *g* is different from 0	1027.97***	50	278	r = .15	[.11, .18]	0.02	7.70	<.0001	… *r* is different from 0	5206.74***
	In-person (*vs.* online)	38	235	g = 0.18	[0.04, 0.32]	0.07	1.51	.1345	… effect of moderator is different from 0		47	223	r = .20	[.13, .27]	0.04	1.47	.1442	… effect of moderator is different from 0	
Type of participant group	Students (intercept)	27	173	g = 0.12	[0.01, 0.23]	0.05	2.26	.0273	… *g* is different from 0	1084.48***	51	244	r = .13	[.09. .18]	0.02	6.52	< .0001	… *r* is different from 0	5607.33***
	Community (*vs.* students)	43	198	g = 0.15	[0.00, 0.29]	0.07	0.34	.7336	… *g* is different from 0		54	284	r = .21	[.15, .27]	0.03	2.33	.0215	… effect of moderator is different from 0	
Outcome variable	Creativity (*vs.* zero)									1051.08***		12	r = .14	[.09, .19]	0.03	5.19	<.0001	… *r* is different from 0	4984.25***
	Empathy (*vs.* zero)		106	g = 0.16	[0.07, 0.24]	0.04	3.81	.0003	… *g* is different from 0			240	r = .11	[.08, .15]	0.02	6.121	<.0001	… *r* is different from 0	
	General cognitive abilities (*vs.* zero)											67	r = .26	[.219, .32]	0.03	7.91	<.0001	… *r* is different from 0	

		Meta-Analysis 1								Meta-Analysis 2							
				Effect Size Estimate								Effect Size Estimate					
		Number of Studies	Number of ES	Parameter	95% *CI*	*SE*	*t*	*p*	*p* informs whether …	*Q*	Number of Studies	Number of ES	Parameter	95% *CI*	*SE*	*t*	*p*	*p* informs whether …	*Q*
	Knowledge (*vs.* zero)		44	*g* = 0.11	[−0.10, 0.32]	0.10	1.07	.2907	… *g* is different from 0										
	Mentalizing (*vs.* zero)		80	*g* = 0.15	[0.03, 0.27]	0.06	2.53	.0141	… *g* is different from 0			35	*r* = .09	[.04, .13]	0.02	3.85	.0002	… *r* is different from 0	
	Moral cognition (*vs.* zero)		79	*g* = 0.08	[-0.03, 0.18]	0.05	1.45	.1518	… *g* is different from 0			74	*r* = .08	[-.01, .17]	0.05	1.78	.0774	… *r* is different from 0	
	Outgroup judgments (*vs.* zero)		52	*g* = 0.14	[-0.07, 0.35]	0.11	1.30	.1982	… *g* is different from 0			25	*r* = .10	[.05, .15]	0.03	3.87	.0002	… *r* is different from 0	
	Verbal abilities (*vs.* zero)											50	*r* = .43	[.35, .50]	0.04	11.40	<.0001	… *r* is different from 0	
	Other thinking processes (*vs.* zero)		9	*g* = 0.25	[-0.07, 0.57]	0.16	1.56	.1245	… *g* is different from 0			53	*r* = .15	[.10, .20]	0.03	5.87	<.0001	… *r* is different from 0	
Type of comparison group	Reading nonfiction (intercept)		269	*g* = 0.10	[0.02, 0.18]	0.04	2.45	.0170	… *g* is different from 0	1059.6830***									
	Reading nothing (*vs.* reading nonfiction)		70	*g* = 0.23	[0.14, 0.32]	0.05	2.93	.0047	… effect of moderator is different from 0										
	Watching fiction (*vs.* reading nonfiction)		14	*g* = 0.31	[0.16, 0.44]	0.07	2.88	.0053	… effect of moderator is different from 0										
Fictionality of print exposure	Nonfiction (intercept)											158	*r* = .09	[.06, .12]	0.02	5.85	< .0001	… *r* is different from 0	6236.28***
	Fiction (*vs.* nonfiction)											559	*r* = .16	[.14, .19]	0.01	5.51	< .0001	… effect of moderator is	

		Meta-Analysis 1									Meta-Analysis 2							
		Number of Studies	Number of ES	Effect Size Estimate					*p* informs whether …	*Q*	Number of Studies	Number of ES	Effect Size Estimate				*p* informs whether …	*Q*
				Parameter	95% *CI*	*SE*	*t*	*p*					Parameter	95% *CI*	*SE*	*t*	*p*		
																		different from 0	
Study design	Correlational											485	*r* = .17	[.14, .21]	0.02	10.11	<.0001	… *r* is different from 0	5783.35***
	Experimental (*vs.* correlational)											73	*r* = .10	[.05, .15]	0.03	-2.70	0.0079	… effect of moderator is different from 0	
Continuous Moderators																			
Participant age		57	326	*b* = −0.01	[−0.01, 0.00]	0	-1.17	.2464	… effect of moderator is different from 0	940.39***	92	481	*b* = −0.00	[−0.01, 0.00]	0	-0.59	.5547	… effect of moderator is different from 0	4567.90***
Percentage of female participants		63	330	*b* = -0.00	[−0.01, 0.00]	0	-0.76	.4476	… effect of moderator is different from 0	1004.26***	103	533	*b* = -0.00	[-0.01, -0.00]	0	-3.49	.0007	… effect of moderator is different from 0	5571.88***
Length of reading assignment		53	315	*b* = -0.00	[−0.00,0.00]	0	-0.99	.3276	… effect of moderator is different from 0	941.09***									

Note. ES = effect sizes. Number of ES but not of studies is reported for variables varying within- and between-study.

Moderator Analyses

A series of regression analyses, parallel to that of Meta-Analysis 1, tested whether the significant amount of heterogeneity can be accounted for by one or several moderator variables. Indeed, meta-regressions yielded several significant moderations. Specifically, type of assessment measure significantly moderated the overall meta-analytic effect – effects were smaller when the assessment of either print exposure or the outcome was based on self-report, r = .13, than when the assessment of neither print exposure nor the outcome was based on self-report, r = .19 (p = .02), whereas effects did not differ significantly between the two constellations in which either both or none of the variables were assessed based on self-report, r = .17 (p = .73). Publication status emerged as a further moderator as the meta-analytic effect was significantly greater for published, r = .18, than for unpublished studies, r = .11 (p = .04). The overall meta-analytic effect was also moderated by type of participant group; effects observed in community samples, r = .21, were significantly greater than those observed in student samples, r = .13 (p = .02).

Regarding the outcome variables, all effects except those associated with moral cognition, r = .08 (p = .08), differed significantly from zero (rs > .08). Pairwise contrasts (p_{crit} = .00179) showed that effects linked with verbal abilities were significantly greater than those associated with all other outcome variables (ps < .0001). In addition, general cognitive abilities were related to significantly smaller effects than verbal abilities, but to greater effects than empathy, mentalizing, and outgroup judgments (ps < .0004). The remaining contrasts were not significant (ps > .013).

Beyond this, fictionality acted as a significant moderator variable since lifetime exposure to written fiction, r = .16, was linked with greater effects than lifetime exposure to

written nonfiction[3], $r = .09$ ($p < .0001$). Also, the overall meta-analytic effect was found to be moderated by study design in that correlational designs, $r = .18$, were linked with significantly greater effects than experimental designs, $r = .10$ ($p = .008$). Finally, percentage of female participants exerted a significant moderating impact, meaning that the higher the percentage of female participants within a study sample, the smaller the overall meta-analytic effect, $b = -0.00$ ($p < .001$). No other moderations reached significance ($ps > .13$).

Publication Bias

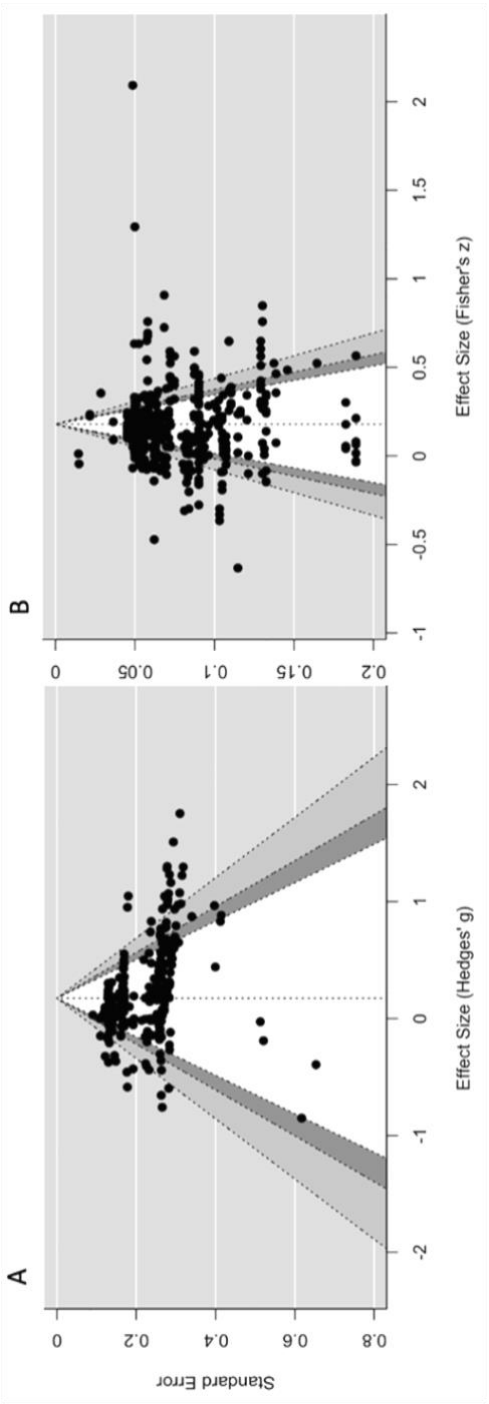
Paralleling Meta-Analysis 1, publication bias analyses included published work only. Visual inspection of the funnel plot suggested some extent of asymmetry since there were more data points the right than to the left of the mean effect (see Figure 2B). Again, the slope of Egger's regression test for funnel plot asymmetry was positive but not significant, $b = 1.33$, $SE = 2.31$, $t(83)$, $p = .57$, reflecting that the precision of the observed effect was not significantly associated with the magnitude of the effect.

Turning to *p*-curve analysis (see Figure 3B), the *p*-curve was significantly right-skewed, $Z = -41.16$, $p < .001$ (full *p*-curve) and $Z = -39.57$, $p < .001$ (half *p*-curve). Hence, the included studies can be assumed to have evidential value in support of the proposal that lifetime exposure to print fiction is correlated with cognitive benefits. Furthermore, the *p*-curve did not reveal evidential inadequacy (i.e., flatter than 33% power), $Z = 27.48$, $p > .999$ (full *p*-curve), and $Z = 36.53$, $p > .999$ (half *p*-curve). The estimated power of the *p*-curve analysis was 98%.

[3] For this particular meta-regression model, a further 158 effects sizes were included that reflected nonfiction exposure. These effect sizes were not used in any other analyses.

Figure 2

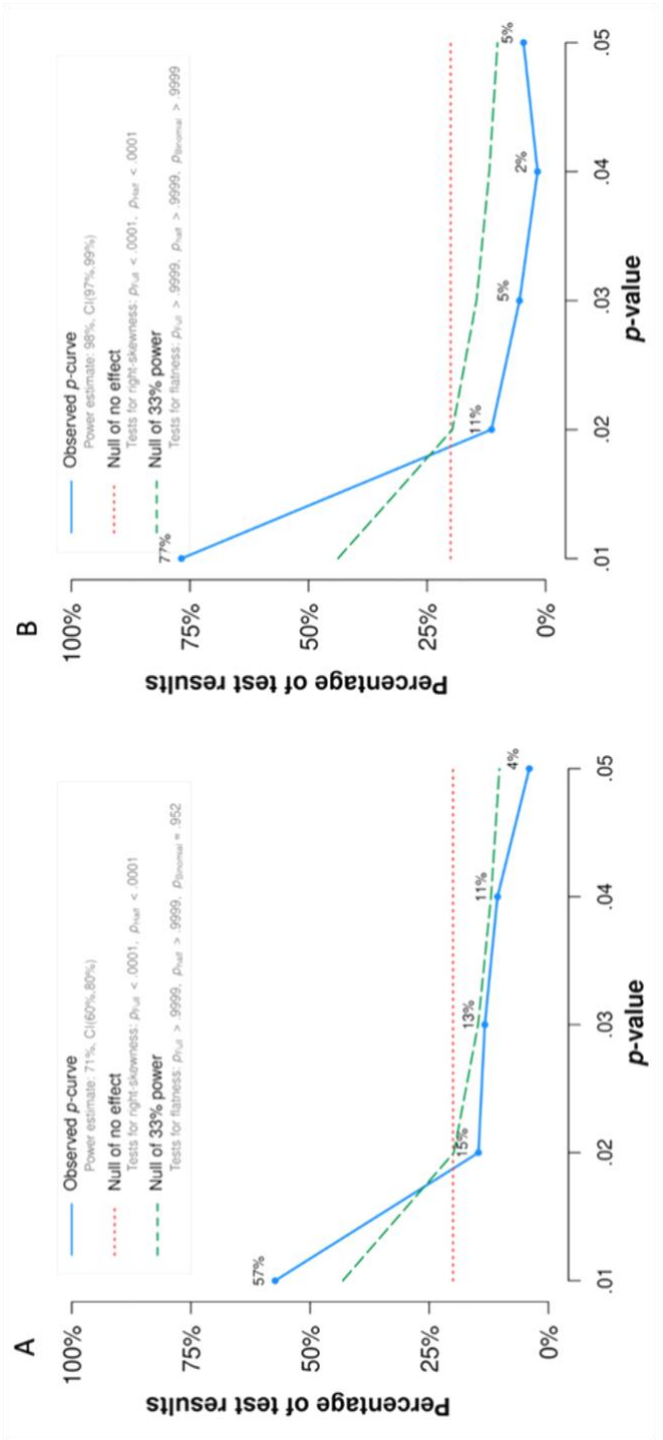
Funnel Plots of the Effect Sizes from Published Studies



Note. Panel A refers to Meta-Analysis 1, panel B to Meta-Analysis 2. Effect sizes on the *x*-axis are plotted against their standard error on the *y*-axis. The dotted vertical line represents the mean meta-analytical effect. Within the funnel shape, the white area reflects the 90% *CI* of the mean effect, the dark grey area reflects the 95% *CI* of the mean effect, and the light grey area reflects the 99% *CI* of the mean effect.

Figure 3

P-Curve Plots Depicting the Distribution of Significant p-Values. Note. Panel A refers to Meta-Analysis 1, panel B to Meta-Analysis 2.



Discussion

In view of ongoing debates about fiction-related cognitive benefits and detriments, the present set of meta-analyses quantitatively synthesized research on cognitive outcomes associated with reading fiction.

Meta-Analysis 1

Meta-Analysis 1 looked at effects of experimental studies involving fiction-reading assignments. First of all, we determined how many experimental studies have investigated cognitive effects of reading fiction. According to our literature search, the relevant number is 70 experiments, which is considerably higher than the body of research (i.e., 14 experiments) summarized by a previous meta-analysis (Dodell-Feder & Tamir, 2018).

Secondly, we were interested in the overall effect of reading fiction on cognition. A three-level random effects meta-analytic model involving 371 effect sizes revealed that compared to control activities, reading fiction led to significant small-sized cognitive benefits, g = 0.14, 95% *CI* [0.06, 0.21], p = .0004. Sensitivity analyses, including checks of standardized residual values and leave-one-out procedures, endorsed the robustness of this effect.

Thirdly, we attempted to identify variables that modulate the overall meta-analytic effect. Two of the variables under investigation were found to exert a moderating influence, namely type of comparison group and outcome variable.

Greater effects emerged when reading fiction was compared with either watching fiction or reading nothing, than when reading fiction was compared with reading nonfiction. This finding touches on the question of whether effects are driven by reading in general, fiction consumption in general, or the combination of reading with consuming fiction (i.e., reading specifically fiction). The current pattern of results suggests that fiction consumption in terms of watching fiction does not have a share in the effect, whereas the activity of reading does account for a portion of the effect. Beyond this, reading particularly fiction was found to make a distinct contribution. This pattern goes against the assumption that consuming

narratives in general, independent of medium (e.g., watching *vs*. reading) and fictionality (i.e., applying to fictional as well as nonfictional narratives), leads to benefits for social cognition. Nevertheless, due to the low number of experiments involving watching fiction, the findings of this meta-regression should be interpreted with caution.

Effects of reading fiction turned out to vary by outcome variable. Significant positive effects emerged for empathy and mentalizing. In contrast, effects on all other outcomes (in particular, knowledge, moral cognition, outgroup judgments, and other thinking processes) did not differ from zero. Additionally, there were no significant differences between any of the outcomes. In sum, only empathy and mentalizing seem to benefit from experimental interventions of reading fiction, and even then, only to a small extent.

Although the overall meta-analytic effect was not significantly greater for published than for unpublished studies ($p = .05$), the latter effect did not differ significantly from zero, $g = 0.04$, 95% *CI* [-0.11, 0.18] (see Table 2). This result indicates that the current pool of studies may be limited by publication bias. However, the remaining publication bias analyses suggested otherwise: according to Egger's test, the funnel plot was not significantly asymmetric and a *p*-curve analysis supported the evidential value of the present body of research. Given that *p*-curve analysis is thought to detect *p*-hacking, it is not a generic tool for assessing publication bias. The file drawer problem (Rosenthal, 1979) exemplifies sources of publication bias not targeted by *p*-curve analysis. Taken together, our analyses do not suggest a publication bias in terms of *p*-hacking, but do not rule out other causes of publication bias, such as the file drawer problem.

The aggregate effect was stable across participant characteristics including gender, age, and type of participant group (students *vs.* community). So, despite the fact that these characteristics are associated with different reading habits or incentives for research participation, respectively, they did not have a significant impact on the synthesized effect. Perhaps reading assignments like the ones used within the experiments have the power to

overrule their influence. Likewise, study format (online *vs.* in-person), although typically linked with different levels of participant supervision, did not act as a significant moderator variable. It is possible that attention checks, which are commonly included in online studies (to adjust for lack of in-person monitoring), served their purpose.

Interestingly, the length of reading assignments did not emerge as a significant moderator. This result does not support the assumption that reading fiction affects cognition via practicing related skills, since such a relation would have elicited a positive relationship. The lack of an effect of reading length could be explained by mechanisms independent of systematic practicing, for instance priming (see Lenhart & Richter, 2022); reading fiction may merely 'warm up' participants' cognitive skills without a lasting transfer effect for real life. It should however be noted that the word count of most reading assignments varied between 1,500 and 6,000 words, which is rather short and within a narrow range that does not provide much room for substantial associations with cognitive outcomes. Moreover, there seems to be consensus that a single reading task involving approximately 6,000 words falls short of the length required for longer lasting effects on cognition (Eekhof et al., 2022; Mar, 2018a, 2018b). Hence, the present body of research does not provide a rigorous test of the assumption that reading fiction leads to genuine changes in cognitive outcomes.

Furthermore, the current meta-analysis is limited in that it may underestimate the true sizes of effects. This is because we only considered post-assessments even though some of the included experiments reported pre-assessments. This analytical approach was taken to replicate that of Dodell-Feder and Tamir (2018) who "used raw unadjusted means and standard deviations, and compared only postreading between-groups scores" (p. 1715). We acknowledge that pre-assessments could have been used to calculate more accurate effect sizes for pre-post designs.

Meta-Analysis 2

Meta-Analysis 2 pooled effects of studies in which lifetime exposure to written fiction was correlated with cognitive skills. Similar to Meta-Analysis 1, our first research question addressed the number of respective studies. Our literature search yielded 114 studies, which exceeded the volume of research (i.e., 30 studies) synthesized by a previous meta-analysis (Mumper & Gerrig, 2017).

The second research question affected the overall correlation of lifetime fiction reading with cognition. According to a three-level random effects meta-analytic model involving 559 effect sizes, there is a significant small-sized positive relationship between lifetime fiction reading volume and cognitive abilities, $r = .16$, 95% *CI* [.13, .19], $p < .0001$. Adopting the same sensitivity analyses as in Meta-Analysis 1, this effect turned out to be robust, so not the result of individual outliers at either the effect size or the study level.

To answer the third research question, we sought to discern moderators of the aggregate meta-analytic effect. Indeed, several of the variables under investigation were found to play a moderating role.

The scenario in which either the outcome or print exposure was assessed via self-report was linked with a lower correlation coefficient than the scenario in which none of these variables was assessed via self-report, but the association did not differ when either both or none of the variables were assessed based on self-report. This pattern goes against the assumption that effects are driven by studies that rely on self-report measures. Among the outcome variables, all correlations but the one for moral cognition were significantly greater than zero. Further, lifetime exposure to written fiction was most strongly related to verbal abilities, next strongly linked with general cognitive abilities, and third strongly linked with empathy, mentalizing, and outgroup judgments. Due to the correlational nature of the studies under investigation, this pattern of results alone does not suggest that improved cognitive outcomes (apart from moral cognition, which was not significantly linked with fiction reading volume) are caused by greater amounts of reading fiction. Instead, individuals with high

cognitive skills could be more strongly attracted to written fictions than individuals with comparatively lower skills. Yet another explanation for the positive correlation between lifetime exposure to written fiction and cognitive skills would be the existence of third-variables causing increases of both frequency of reading fiction and cognitive outcomes.

A further meta-regression is relevant to the question of whether effects on cognition are specific to reading fiction or apply to reading in general. Lifetime exposure to written fiction was linked with a stronger correlation than lifetime exposure to written nonfiction, although both correlations were significantly positive. In line with findings of Meta-Analysis 1, the activity of reading seems to contribute to the overall effect. Additionally, reading fiction in particular appears to be associated with distinct benefits. In terms of participant characteristics, the overall meta-analytic effect was larger in community samples, who often receive monetary compensation in exchange for participation, than in student samples, who are typically reimbursed through course credits. Yet, these sample types differed on other characteristics as well, namely age (M = 29.13, SD = 12.63 *vs.* M = 20.75, SD = 2.16), study format (64.1% online *vs* 34.2% online), and percentage of female participants (M = 58.75%, SD = 10.43 *vs* M = 73.18%, SD = 11.24). Of these, only the percentage of females proved to moderate the aggregate meta-analytic effect: the higher the proportion of females, the lower the association between lifetime fiction reading and cognition. Possibly, males benefit more from reading fiction than females. Given that males tend to read less fiction than females, there may be greater room for fiction-based improvements among them. In addition, since the percentage of female participants differed between the two sample types (community *vs.* students), it is possible that the difference between sample types was confounded by different incentives or gender balance.

Finally, publication status emerged as a significant moderator – the aggregate correlation was significantly greater for published than for unpublished studies, with significant overall effects at both levels of this moderator variable. Even though this finding

suggests the presence of a publication bias, results of the remaining publication bias analyses were contradictory: the funnel plot was not found to be significantly asymmetric; additionally, a *p*-curve analysis supported the evidential value of the current pool of studies. So, just as in Meta-Analysis 1, the tests do not indicate publication bias in the sense of *p*-hacking, but do not exclude other reasons for publication bias such as the file drawer problem.

Limitations of the Two Pools of Studies

The studies included in both meta-analyses are limited with regards to the cognitive outcomes under investigation. On the one hand, our meta-analyses suggest that reading fiction is associated with small-sized benefits for cognition. On the other hand, they offer only weak support for the assumption that reading fiction is *not* linked with adverse effects on cognition as discussed in the Introduction. This is because in Meta-Analysis 1, only 19.68% of effect sizes concerned any undesirable types of cognition, for example negative perceptions of robots (Mara & Appel, 2015), sexual prejudice (Orellana, 2019, Experiment 1), or pejorative attitudes towards autistic individuals (Stern, 2020; these effect sizes were reverse coded, so that negative effect sizes indicate cognitive detriments). Importantly, a number of experiments on acquiring misinformation (but, importantly, also accurate information) from fiction were excluded since they did not include a comparison condition not involving reading fiction (e.g., Appel & Richter, 2007; Butler et al., 2012; Eslick et al., 2011; Fazio et al., 2013; Gerrig & Prentice, 1991; Marsh & Fazio, 2006; Marsh et al., 2003; Prentice et al., 1997; Rapp et al., 2014; Wheeler et al., 1999). In Meta-Analysis 2, only 11.09% of effect sizes concerned undesirable types of cognition, for instance, unrealistic perceptions of romantic relationships (Stern et al., 2019), egocentric bias (Castano et al., 2020), or psychological essentialism (Castano et al., 2021). Like all of the cognitive outcomes under investigation, however, these variables include a broad range of cognitions such as attitudes (e.g., sexual prejudice, pejorative attitudes toward autistic individuals) and cognitive biases (e.g., egocentric bias, psychological essentialism). In sum, further research (with appropriate control conditions

wherever applicable) is needed to assess whether reading fiction is associated with unwanted cognitive outcomes.

A further limitation concerns underlying textual mechanisms. Regarding Meta-Analysis 1, the fiction reading stimuli used within experiments were typically narrative fictions, in which effects of narrativity and fictionality are conflated. Therefore, we do not know whether effects apply to non-narrative fictions such as certain types of poetry, even if the latter are in reality less often read than narrative fictions (Jackson, 2008). In addition, researchers generally did not describe the literariness of the reading material in sufficient detail or did not mention the literariness of their reading stimuli at all. Consequently, like Dodell-Feder and Tamir (2018), we were not able to explicitly consider literariness as a moderator variable in our meta-analysis. Relatedly, study authors usually did not address the extent to which their reading stimuli represent the textual genre of interest. Thus, it is impossible to know whether the current results pertain to the reading stimuli used within experiments or expand to the entire fiction genre (see Liberman, 2013). Similarly, in Meta-Analysis 2, the measures of fiction exposure under investigation typically assess the frequency with which narrative fictions are read, but do not touch on non-narrative fictions. So again, we do not know whether the overall meta-analytic effect generalizes to non-narrative fictions. Even though approximately 9% of studies collected exposure to especially literary fictions, the vast majority did not address literariness. In sum, we could not take into account literariness as a moderator.

There remain open questions regarding the underlying textual mechanisms of fiction reading effects. In other words, whether the effects summarized here are due to the content of fiction (e.g., because works of fiction contain predominantly accurate or inaccurate information that is acquired during reading) or due to engagement with fictions (e.g., because reading fictions draws on cognitive processes that are strengthened as a result; see also Mar, 2018a, who distinguishes between a process- and a content-based route to learning from

narratives). The original research reports included in Meta-Analysis 1 do not typically provide an in-depth analysis of either the reading materials' content or the cognitive processes engaged in during reading. For the original work synthesized in Meta-Analysis 2, such an analysis seems almost impossible given that those studies are about associations of accumulated reading experiences with cognitive outcomes assessed offline. Hence, neither of the present meta-analyses can inform whether the presently observed cognitive benefits and disadvantages are related to the content of fictions or to the cognitive processes engaged in during reading.

Furthermore, the quality of the evidence was mixed in both Meta-Analyses: In Meta-Analysis 1, 44% of effect sizes harbored high risk of bias in that they assessed the outcome via self-report; furthermore, sample sizes lacked sufficient power to identify an effect of the size observed – fiction reading groups had a median sample size of 51, control groups a median sample size of 48. According to a post-hoc power analysis (using G*Power; Faul et al., 2007), this sample size had a power of 11% to detect an effect of the size observed (i.e., $d = 0.14$) at the standard 5% significance level in a two-tailed *t*-test for independent samples. In Meta-Analysis 2, both print exposure and the outcome were assessed via self-report for only 6.46% of effect sizes, yet 49.37% assessed either print exposure or the outcome via self-report, so that neither print exposure nor the outcome were assessed via self-report for just approximately half of the effect sizes. Beyond this, sample sizes suffered from a similar power issue as sample sizes of Meta-Analysis 1. Studies worked with a median sample size of 194. A post-hoc power analysis (again using G*Power) revealed that this sample size had insufficient power (59%) to detect a small-sized correlation coefficient of $r = .16$, as observed here, at the standard 5% significance level (two-tailed). Also, the median sample size is below the target of $n = 250$ at which correlation coefficients have been found to stabilize (Schönbrodt & Perugini, 2013).

Similarities and Differences with Previous Syntheses

The aggregate effect of reading fiction in Meta-Analysis 1 (g = 0.14) was almost identical to that reported by Dodell-Feder and Tamir's (2018) previous synthesis (g = 0.15). Similarly, the total effect of lifetime fiction reading in Meta-Analysis 2 (r = .16) was comparable to observations in Mumper and Gerrig's (2017) previous synthesis. These authors did not calculate an overall effect for social cognition, but separate scores for empathy, r = .07, and mentalizing, r = .21. The present synthesized effect falls in between these coefficients. Differences between the current and both previous meta-analyses pertain to moderators and publication bias. The fact that the earlier syntheses did not yield significant outputs in either regard whereas the present ones did – at least for some analyses – may be related to differences in power. More precisely, Meta-Analysis 1 involved 5 times the number of studies included in Dodell-Feder and Tamir's (2018) work, whilst Meta-Analysis 2 covered 3.8 times the number of investigations involved in Mumper and Gerrig's (2017) synthesis; hence, power was much improved for the present meta-analyses.

Summary and Conclusions

Having discussed the findings of each meta-analysis in turn, we close with a comparative summary. Both meta-analyses yielded small-sized aggregate effects, suggesting that short fiction reading assignments cause small cognitive benefits and that lifetime exposure to written fiction is related to small cognitive enhancements, respectively. Overall effects of both meta-analyses were robust, so were neither the result of individual outliers nor driven by assessments based on self-report. Given that undesirable outcomes were rarely investigated, the two meta-analyses provide only minimal evidence for the assumption that reading fiction is not linked with cognitive detriments. Thus, there is uncertainty about any negative outcomes that may accrue to those who make a steady *diet* of reading fiction, and so the possible existence of negative effects remains a crucial question. Results of both meta-analyses also converged regarding the role of reading fiction for the aggregate effects: Both

projects showed that reading in general makes a significant contribution, with fiction reading having a particular impact.

In terms of the outcomes associated with fiction exposure, a significant effect of reading fiction on moral cognition did not arise in Meta-Analysis 1, and was even more strikingly missing in Meta-Analysis 2, in which moral cognition was the only cognitive outcome that was not significantly associated with reading fiction. This pattern conflicts with some prevailing philosophical theories of morality such as moral sentimentalism (Kauppinen, 2022) which posit that most of the content (i.e., what is reasoned about or decided upon) in moral cognition is essentially emotional in nature. This philosophical position has been supported by neuroscientific evidence as well (e.g., Young & Koenigs, 2007). Given that empathy is defined by an emotional resonance or connection between an observer and an object of observation (typically another person), it is noteworthy that reading fiction would not have some pronounced effect on moral reasoning.

Beyond this, publication bias is a concern for the two meta-analyses. This bias does not seem to be associated with *p*-hacking, but to reflect the file drawer problem (Rothstein, 2008). Further limitations shared by both meta-analyses include risk of bias linked with self-report assessments, underpowered sample sizes, and insufficient consideration of narrativity and literariness as potential mechanisms.

The patterns of results for the two meta-analyses differed in two major ways. Firstly, the percentage of female participants was a significant moderator of correlational (Meta-Analysis 2), but not of experimental effects (Meta-Analysis 1). It is possible that males have a greater potential for fiction-related cognitive benefits since they have a lower preference for written fictions than females (Hu et al., 2023). Alternatively, the gender effect seen in Meta-Analysis 2 might reflect confounds, due to differences between student and community samples, such as varying incentives for research participation. In any case, the fact that participant characteristics did not moderate effects in Meta-Analysis 1 suggests that short

fiction-reading assignments can overrule the influence of reader attributes. Secondly, the meta-analyses yielded diverging results regarding the cognitive outcomes associated with reading fiction. Experimental effects of reading fiction were significant for indicators of social cognition only (i.e., empathy and mentalizing), but correlations with lifetime fiction exposure were significant for all cognitive outcomes except for moral cognition. In fact, effects of empathy and mentalizing were surpassed by those of verbal and general cognitive abilities, and were not stronger than effects of remaining outcomes.

If we assume a causal impact of reading fiction across both meta-analyses, this could indicate that short fiction-reading assignments merely prime social cognitive skills, and that these priming effects consolidate over time without growing in size. The other cognitive outcomes that correlated with lifetime exposure to print fiction may not be immediately primed during reading, at least not to a measurable extent, but may still accumulate over time. Alternatively, the pattern could be interpreted as evidence against a causal impact of reading fiction: If effects for verbal and general cognitive abilities show up only in correlational studies, which cannot confirm causal relationships, but do not become evident in experiments, this may suggest that reading fiction does not cause sustainable cognitive benefits. In that case, the aggregate effect obtained in Meta-Analysis 1 may reflect a transient priming response and the overall effect yielded in Meta-Analysis 2 may reflect differences in fiction-reading preferences between people high in verbal and/or general cognitive abilities. Finally, third-variables, such as education level, could underlie the association between lifelong fiction reading and cognition. This explanation is not unlikely given the very small effect sizes in Meta-Analysis 2 (and in the earlier meta-analysis by Mumper & Gerrig, 2017), and the fact that we relied on raw correlations (not partial correlations controlling for the impact of third variables). Lifetime exposure to written fiction might then not be the cause of the cognitive benefits observed.

To conclude, we have presented a comprehensive and pre-registered set of meta-analyses that provide robust evidence for a small positive relationship between reading fiction and cognitive benefits. Whether this relation represents a sustainable change caused by reading fiction remains unknown. Further research needs to: i) examine the underlying textual mechanisms and potential for adverse outcomes of reading fiction; ii) implement stronger manipulations of reading volume; iii) assess variables based on behavioral performance wherever possible; iv) recruit well-powered samples; and v) consider publication of results independent of statistical significance, an invitation also directed to publication outlets.

Constraints on Generality

For most studies from both meta-analyses, samples were younger adults between approximately 20 and 34 years of age, with the percentage of female participants varying between approximately 53 and 77%. Hence, results are representative of younger adult groups with a rather high percentage of females, whereas children/adolescents, older adults, and also males, are underrepresented. Future research with these understudied groups would be desirable.

References

References marked with † indicate studies included in Meta-Analysis 1. References marked with ‡ indicate studies included in Meta-Analysis 2.

‡ Allen, L. J. S., Cipielewski, J., & Stanovich, K. E. (1992). Multiple indicators of children's reading habits and attitudes: Construct validity and cognitive correlates. *Journal of Educational Psychology, 84*(4), 489–503. https://doi.org/10.1037/0022-0663.84.4.489

Anuar, A. F. binti, & Wan, A. J. binti A. C. (2021). Finding comfort: A study of reading habits among Malaysian University students during Covid-19 pandemic. *International Journal of Academic Research in Business & Social Sciences, 11*(11), 894-909. https://doi.org/10.6007/ijarbss/v11-i11/11583Appel, M., Hanauer, D. A., Hoeken, H., Van Krieken, K., Richter, T., & Sanders, J. (2021). The psychological and social

effects of literariness: Formal features and paratextual information. In D. Kuiken & A. M. Jacobs (Eds.), *Handbook of empirical literary studies* (pp. 177–202). De Gruyter.

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

‡ Arnold, J. J., Castro-Schilo, L., Zerkle, S. A., & Rao, L. (2019). Print exposure predicts pronoun comprehension strategies in children. *Journal of Child Language*, *46*(5), 863–893. https://doi.org/10.1017/s0305000919000102

‡ Arnold, J. J., Strangmann, I. M., Hwang, H., Zerkle, S. A., & Nappa, R. (2018). Linguistic experience affects pronoun interpretation. *Journal of Memory and Language*, *102*, 41–54. https://doi.org/10.1016/j.jml.2018.05.002

Bakhtiari, A., Bjørke, A. B., Larsson, P. G., Olsen, K. B., Nævra, M. C. J., Taubøll, E., Heuser, K., & Østby, Y. (2022). Episodic memory dysfunction and effective connectivity in adult patients with newly diagnosed nonlesional temporal lobe epilepsy. *Frontiers in Neurology*, *13*, 774532. https://doi.org/10.3389/fneur.2022.774532

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

Barnes, J. L. (2018). Imaginary engagement, real-world effects: Fiction, emotion, and social cognition. *Review of General Psychology*, *22*(2), 125–134. https://doi.org/10.1037/gpr0000124

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

Bastian, B., & Haslam, N. (2006). Psychological essentialism and stereotype endorsement. *Journal of Experimental Social Psychology*, *42*(2), 228–235. https://doi.org/10.1016/j.jesp.2005.03.003

Best, J. (2020). Reading literary fiction: More empathy, but at what possible cost? *North American Journal of Psychology*, *22*(2), 269-288.

Best, J. (2021). To teach and delight: The varieties of learning from fiction. *Review of General Psychology*, *25*(1), 27–43. https://doi.org/10.1177/1089268020977173

‡ Bischoff, T. (2013). Interior music: An examination of the sociocognitive abilities of fiction writers [Unpublished doctoral dissertation]. University of Toronto.

‡ Black, E. (1959). *A study of science reasoning abilities of science fiction readers*. [Unpublished master's thesis]. Atlanta University.

‡ Black, J. (2018). *The story of moral character: Morality, narrative, and the value of striving* [Unpublished doctoral dissertation]. The University of Oklahoma.

† ‡ Black, J. M., & Barnes, J. L. (2015a). 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. M., & Barnes, J. L. (2015b). 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. M., & Barnes, J. L. (2021a). Fiction and morality: Investigating the associations between reading exposure, empathy, morality, and moral judgment. *Psychology of Popular Media*, *10*(2), 149–164. https://doi.org/10.1037/ppm0000281

‡ Black, J. M., & Barnes, J. L. (2021b). Pushing the boundaries of reality: Science fiction, creativity, and the moral imagination. *Psychology of Aesthetics, Creativity, and the Arts*, *15*(2), 284–294. https://doi.org/10.1037/aca0000281

‡ Black, J. M., & Barnes, J. L. (2021c). Does writing promote social cognition? The role of fictionality and social content. *Psychology of Aesthetics, Creativity, and the Arts*. https://doi.org/10.1037/aca0000438

Black, J. M., 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. Jacobs (Eds.), *Handbook of empirical literary studies* (pp. 229-250). De Gruyter.

‡ Black, J. M., Capps, S. C., & Barnes, J. L. (2018). Fiction, genre exposure, and moral reality. *Psychology of Aesthetics, Creativity, and the Arts*, *12*(3), 328–340. https://doi.org/10.1037/aca0000116

Black, J. E., & Reynolds, W. M. (2016). Development, reliability, and validity of the Moral Identity Questionnaire. *Personality and Individual Differences*, *97*, 120–129. https://doi.org/10.1016/j.paid.2016.03.041

Borenstein, M. (Ed.). (2009). *Introduction to meta-analysis*. John Wiley & Sons.

Bradshaw, T., & Nichols, B. (2004). *Reading at risk: A survey of literary reading in America* (ED484208). ERIC. https://files.eric.ed.gov/fulltext/ED484208.pdf

Brown, J. I., Bennett, J. M., & Hanna, G. S. (1993). *The Nelson-Denny reading test: Vocabulary, comprehension, reading rate. Forms E and F*. Riverside

† Browning, E., & Hohenstein, J. (2015). The use of narrative to promote primary school children's understanding of evolution. *Education 3-13*, *43*(5), 530–547. https://doi.org/10.1080/03004279.2013.837943

‡ Brysbaert, M., Sui, L., Dirix, N., & Hintz, F. (2020). Dutch author recognition test. *Journal of Cognition*, *3*(1), 6. https://doi.org/10.5334/joc.95

‡ Brysbaert, M., & Vantieghem, A. (2022). *No correlation between articulation speed and silent reading rate when adults read short texts* [Unpublished article]. Ghent University.

Burke, M., Kuzmicova, A., Mangen, A., & Schilhab, T. (2016). Empathy at the confluence of neuroscience and empirical literary studies. *Scientific Study of Literature, 6*(1), 6–41. https://doi.org/10.1075/ssol.6.1.03bur

Butler, A., Dennis, N., & Marsh, E. (2012). Inferring facts from fiction: Reading correct and incorrect information affects memory for related information. *Memory*, *20*(5), 487–498. https://doi.org/10.1080/09658211.2012.682067

‡ Castano, E., Martingano, A. J., & Perconti, P. (2020). The effect of exposure to fiction on attributional complexity, egocentric bias and accuracy in social perception. *PLOS ONE*, *15*(5), e0233378. https://doi.org/10.1371/journal.pone.0233378

‡ Castano, E., Paladino, M. P., Cadwell, O. G., Cuccio, V., & Perconti, P. (2021). Exposure to literary fiction is associated with lower psychological essentialism. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.662940

Cattell, R. B., & Scheier, I. H. (1961). *The meaning and measurement of neuroticism and anxiety*. Ronald.

† Cervetti, G. N., Bravo, M. A., Hiebert, E. H., Pearson, P. D., & Jaynes, C. A. (2009). Text genre and science content: Ease of reading, comprehension, and reader preference. *Reading Psychology*, *30*(6), 487–511. https://doi.org/10.1080/02702710902733550

† Chlebuch, N., Goldstein, T. R., & Weisberg, D. S. (2020). Fact or fiction? Clarifying the relationship between reading and the improvement of social skills. *Scientific Study of Literature, 10*(2), 167-192. https://doi.org/10.1075/ssol.20007.chl

Cho, H., Lee, W., Urban, A., Huang, L., & Long, Y. (2021). "I don't want a book that's going to make me sad or stressed out, especially in this day and age": Fiction reading (and healing) in a pandemic. *Proceedings of the Association for Information Science and Technology*, *58*(1), 420–424. https://doi.org/10.1002/pra2.471

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

Consoli, G. (2018). Preliminary steps towards a cognitive theory of fiction and its effects. *Journal of Cultural Cognitive Science, 2*(1–2), 85–100. https://doi.org/10.1007/s41809-018-0019-5

Costanzo, M., & Archer, D. (1993). *The Interpersonal Perception Task-15 (IPT-15): A guide for researchers and teachers*. Unpublished Manual.

Cuff, B. M. P., Brown, S. J., Taylor, L., & Howat, D. J. (2016). Empathy: A review of the concept. *Emotion Review, 8*(2), 144–153. https://doi.org/10.1177/1754073914558466

† Cunningham, L. A., & Gall, M. D. (1990). The effects of expository and narrative prose on student achievement and attitudes toward textbooks. *Journal of Experimental Education, 58*(3), 165–175. https://doi.org/10.1080/00220973.1990.10806532

‡ Cunningham, A. E., & Stanovich, K. E. (1991). Tracking the unique effects of print exposure in children: Associations with vocabulary, general knowledge, and spelling. *Journal of Educational Psychology, 83*(2), 264–274. https://doi.org/10.1037/0022-0663.83.2.264

Currie, G. (1985). What is fiction? *The Journal of Aesthetics and Art Criticism, 43*(4), 385-392. https://doi.org/10.2307/429900

Currie, G. (2020) *Imagining and knowing: The shape of fiction*. Oxford University Press.

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology, 10*.

Davies, B., Lupton, C., & Gormsen Schmidt, J. (2022). *Reading novels during the Covid-19 pandemic* (1st ed.). Oxford University Press. https://doi.org/10.1093/oso/9780192857682.001.0001

† De Mulder, H., Hakemulder, F., Van Den Berghe, R., Klaassen, F., & Van Berkum, J. (2017a). Effects of exposure to literary narrative fiction. *Scientific Study of Literature, 7*(1), 129–169. https://doi.org/10.1075/ssol.7.1.06dem

‡ De Mulder, H., Hakemulder, F., Klaassen, F., Junge, C., Hoijtink, H., & Van Berkum, J. J. A. (2021). Figuring out what they feel: Exposure to eudaimonic narrative fiction is related to mentalizing ability. *Psychology of Aesthetics, Creativity, and the Arts*, *16*(2), 242–258. https://doi.org/10.1037/aca0000428

‡ De Mulder, H., Hakemulder, F., Van Den Berghe, R., Klaassen, F., & Van Berkum, J. (2017b). Effects of exposure to literary narrative fiction. *Scientific Study of Literature*, *7*(1), 129–169. https://doi.org/10.1075/ssol.7.1.06dem

‡ Diekman, A. B., McDonald, M., & Gardner, W. L. (2000). Love means never having to be careful. *Psychology of Women Quarterly*, *24*, 179-188.

Dippo, C. (2013). *Evaluating the Alternative Uses Test of creativity*. Proceedings of the National Conference, On Undergraduate Research (NCUR) 2013. University of Wisconsin La Crosse.

† Dodell-Feder, D., Liu, G., Amormino, P., Handley, E., & Tamir, D. (2022). *The impact of fiction reading on social outcomes: A 4-week randomized controlled study*. University of Rochester.

Dodell-Feder, D., & Tamir, D. I. (2018). Fiction reading has a small positive impact on social cognition: A meta-analysis. *Journal of Experimental Psychology*, *147*(11), 1713–1727. https://doi.org/10.1037/xge0000395

Dubourg, E., & Baumard, N. (2022). Why and how did narrative fictions evolve? Fictions as entertainment technologies. *Frontiers in Psychology, 13, 786770*. https://doi.org/10.3389/fpsyg.2022.786770

Eco, U. (1994). *Six walks in the fictional woods*. Harvard University Press.

‡ Eekhof, L. S., Van Krieken, K., Sanders, J., & Willems, R. M. (2023). Engagement with narrative characters: The role of social-cognitive abilities and linguistic viewpoint. *Discourse Processes, 60*(6), 411–439. https://doi.org/10.1080/0163853X.2023.2206773

‡ Eekhof, L. S., Van Krieken, K., Sanders, J., & Willems, R. M. (2021). Reading minds, reading stories: Social-cognitive abilities affect the linguistic processing of narrative viewpoint. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.698986

Eekhof, L. S., Van Krieken, K., & Willems, R. M. (2022). Reading about minds: The social-cognitive potential of narratives. *Psychonomic Bulletin & Review*, *29*(5), 1703–1718. https://doi.org/10.3758/s13423-022-02079-z

Effron, D. A., & Conway, P. (2015). When virtue leads to villainy: Advances in research on moral self-licensing. *Current Opinion in Psychology*, *6*, 32–35. https://doi.org/10.1016/j.copsyc.2015.03.017

Eidelson, R. J., & Epstein, N. (1982). Cognition and relationship maladjustment: Development of a measure of dysfunctional relationship beliefs. *Journal of Consulting and Clinical Psychology*, *50*(5), 715–720. https://doi.org/10.1037/0022-006X.50.5.715

Erikson, M., Erikson, M. G., & Punzi, E. (2020). Reading fiction as a learning activity in clinical psychology education: Students' perspectives. *International Journal of Teaching and Learning in Higher Education*, *32*(2), 171-179.

Eslick, A. N., Fazio, L. K., & Marsh, E. J. (2011). Ironic effects of drawing attention to story errors. *Memory*, *19*(2), 184–191. https://doi.org/10.1080/09658211.2010.543908

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fazio, L. K., Barber, S. J., Rajaram, S., Ornstein, P. A., & Marsh, E. J. (2013). Creating illusions of knowledge: Learning errors that contradict prior knowledge. *Journal of Experimental Psychology: General*, *142*(1), 1–5. https://doi.org/10.1037/a0028649

Ferguson, A. M., Cameron, C. D., & Inzlicht, M. (2021). When does empathy feel good? *Current Opinion in Behavioral Sciences*, *39*, 125–129. https://doi.org/10.1016/j.cobeha.2021.03.011

Ferguson, H. J., & Wimmer, L. (2023). A psychological exploration of empathy. In F. Mezzenzana & D. Peluso (Eds.), *Conversations on empathy* (1st ed., pp. 60–77). Routledge. https://doi.org/10.4324/9781003189978-5

Fiske, S. T., & Taylor, S. E. (2013). *Social cognition: From brains to culture* (2nd ed.). SAGE Publications.

Förster, J., Friedman, R. S., & Liberman, N. (2004). Temporal construal effects on abstract and concrete thinking: Consequences for insight and creative cognition. *Journal of Personality and Social Psychology*, *87*(2), 177–189. https://doi.org/10.1037/0022-3514.87.2.177

‡ 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

‡ Fong, K., Mullin, J. B., & Mar, R. A. (2015). How exposure to literary genres relates to attitudes toward gender roles and sexual behavior. *Psychology of Aesthetics, Creativity, and the Arts*, *9*(3), 274–285. https://doi.org/10.1037/a0038864

Fox, J. (1991). *Regression diagnostics: An introduction*. SAGE Publications.

Friend, S. (2012). VIII-Fiction as a genre. *Proceedings of the Aristotelian Society for the Systematic Study of Philosophy*, *112*(2pt2), 179–209. https://doi.org/10.1111/j.1467-9264.2012.00331.x

† Furnham, A. (2001). Remembering stories as a function of the medium of presentation. *Psychological Reports*, *89*(3), 483-486.

George, T., & Wiley, J. (2019). Fixation, flexibility, and forgetting during alternate uses tasks. *Psychology of Aesthetics, Creativity, and the Arts*, *13*(3), 305–313. https://doi.org/10.1037/aca0000173

Gertken, J., & Köppe, T. (2009). Fiktionalität [Fictionality]. In S. Winko, F. Jannidis, & G. Lauer (Eds.), *Grenzen der Literatur. Zu Begriff und Phänomen des Literarischen*

(Revisionen Grundbegriffe des Literaturtheorie) [Frontiers of literature. On the notion and phenomenon of literariness (Revisions fundamental of literary theory)] (Vol. 2, pp. 228–266). De Gruyter.

Gerrig, R. J., & Prentice, D. A. (1991). The representation of fictional information. *Psychological Science, 2*(5), 336–340. https://doi.org/10.1111/j.1467-9280.1991.tb00162.x

‡ Gitau, K. B. (2002). *Influence of the reading of fiction on secondary school performance of English in Mbitini division of Makueni district* [Unpublished master's thesis]. University of Nairobi.

Glick, P., & Fiske, S. T. (1997). Hostile and benevolent sexism: Measuring ambivalent sexist attitudes toward women. *Psychology of Women Quarterly, 21*(1), 119–135. https://doi.org/10.1111/j.1471-6402.1997.tb00104.x

Goffin, K., & Friend, S. (2022). Learning implicit biases from fiction. *The Journal of Aesthetics and Art Criticism*, *80*(2), 129–139. https://doi.org/10.1093/jaac/kpab078

Goldstein, T. R., & Winner, E. (2012). Enhancing empathy and theory of mind. *Journal of Cognition and Development*, *13*(1), 19–37. https://doi.org/10.1080/15248372.2011.573514

Gottfredson, L. S. (1997). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence, 24*(1), 13–23.

‡ Grant, A. K. (2011). *Exploring the relationship between genre preferences and reading skills: Are boys really such poor readers?* [Doctoral dissertation, Wilfried Laurier University]. Library and Archives Canada.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*(6), 1464–1480. https://doi.org/10.1037/0022-3514.74.6.1464

Guilford, J. P. (1967). *The nature of human intelligence.* McGraw-Hill.

Håkansson Eklund, J., & Summer Meranius, M. (2021). Toward a consensus on the nature of empathy: A review of reviews. *Patient Education and Counseling*, *104*(2), 300–307. https://doi.org/10.1016/j.pec.2020.08.022

† Hakemulder, F. (2000). *The moral laboratory: Experiments examining the effects of reading literature on social perception and moral self-concept*. John Benjamins.

† Hakemulder, J. (2008). Imagining what could happen. In S. Zyngier, M. Bortolussi, A. Chesnokova, & J. Auracher (Eds.), *Directions in empirical literary studies* (pp. 139-153). John Benjamins.

‡ Hall, J. G., Goh, J. X., Mast, M. S., & Hagedorn, C. (2016). Individual differences in accurately judging personality from text. *Journal of Personality*, *84*(4), 433–445. https://doi.org/10.1111/jopy.12170

Happé, F., Cook, J. L., & Bird, G. (2017). The structure of social cognition: In(ter)dependence of sociocognitive processes. *Annual Review of Psychology*, *68*(1), 243–267. https://doi.org/10.1146/annurev-psych-010416-044046

Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2022). *Doing meta-analysis with R: A hands-on guide* (1st ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9781003107347

Harrison, A. J., Bradshaw, L. P., Naqvi, N. C., Paff, M. L., & Campbell, J. M. (2017). Development and psychometric evaluation of the Autism Stigma and Knowledge Questionnaire (ASK-Q). *Journal of Autism and Developmental Disorders*, *47*(10), 3281–3295. https://doi.org/10.1007/s10803-017-3242-x

† ‡ Hollis, H. (2022). *The influence of reading fiction upon critical thinking* [Unpublished doctoral dissertation]. University College London.

Hooper, M. (2020, March). *Troubling trends: An international decline in reading attitudes. IEA Compass: Briefs in Education No. 8*. (ED 608237). ERIC. https://files.eric.ed.gov/fulltext/ED608237.pdf

Hopkins, E., & Weisberg, D. S. (2017). The youngest readers' dilemma: A review of children's learning from fictional sources. *Developmental Review, 43*, 48–70. https://doi.org/10.1016/j.dr.2016.11.001

Hu, J., Yan, G., Wen, X., & Wang, Y. (2023). Gender differences in reading medium, time, and text types: Patterns of student reading habits and the relation to reading performance. *Reading and Writing*. https://doi.org/10.1007/s11145-023-10446-y

‡ Jackson, S. E. (2022). Exploring the mutually reinforcing relationship between theory of mind and reading in adult readers. *Journal of Research in Reading*, *45*(2), 189–203. https://doi.org/10.1111/1467-9817.12393

Jackson, V. (2008). Who reads poetry? *PMLA*, *123*(1), 181–187. http://doi.org/10.1632/pmla.2008.123.1.181

Jacobs, A. M., & Willems, R. M. (2018). The fictive brain: Neurocognitive correlates of engagement in literature. *Review of General Psychology*, *22*(2), 147–160. https://doi.org/10.1037/gpr0000106

† Johnson, D. L., Huffman, B. L., & Jasper, D. M. (2014). Changing race boundary perception by reading narrative fiction. *Basic and Applied Social Psychology*, *36*(1), 83–90. https://doi.org/10.1080/01973533.2013.856791

† Johnson, D. L., Jasper, D., Griffin, S., & Huffman, B. L. (2013). Reading narrative fiction reduces Arab-Muslim prejudice and offers a safe haven from intergroup anxiety. *Social Cognition*, *31*(5), 578–598. https://doi.org/10.1521/soco.2013.31.5.578

‡ Johnson, E., & Arnold, J. E. (2021). Individual differences in print exposure predict use of implicit causality in pronoun comprehension and referential prediction. *Frontiers in Psychology*, *12*, 672109. https://doi.org/10.3389/fpsyg.2021.672109

† Justice, L. M., Meier, J. D., & Walpole, S. (2005). Learning new words from storybooks. *Language Speech and Hearing Services in Schools*, *36*(1), 17–32. https://doi.org/10.1044/0161-1461(2005/003)

‡ Kangassalo, M. (2021). *How habits of media consumption relate to scientific literacy: A quantitative and qualitative analysis* [Master's thesis, Tampere University]. Dspace. https://urn.fi/URN:NBN:fi:tuni-202109227211

Kauppinen, Antti, "Moral Sentimentalism", The Stanford Encyclopedia of Philosophy (Spring 2022 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2022/entries/moral-sentimentalism/>.

† Kelly, C. (2008). *Comparison of textbook passages, nonfiction trade book passages and fiction trade book passages as instructional tools for learning science* (Publication No. 3348505) [Doctoral dissertation, The Florida State University]. ProQuest Dissertations.

† Kidd, D. G., & Castano, E. (2013). Reading literary fiction improves theory of mind. *Science, 342*(6156), 377–380. https://doi.org/10.1126/science.1239918

† Kidd, D. G., & Castano, E. (2017a). Panero et al. (2016): Failure to replicate methods caused the failure to replicate results. *Journal of Personality and Social Psychology*, *112*(3), e1–e4. https://doi.org/10.1037/pspa0000072

‡ Kidd, D. G., & Castano, E. (2017b). Different stories: How levels of familiarity with literary and genre fiction relate to mentalizing. *Psychology of Aesthetics, Creativity, and the Arts*, *11*(4), 474–486. https://doi.org/10.1037/aca0000069

† Kidd, D. G., & 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. G., 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

† Kilaru, A. S., Perrone, J., Auriemma, C. L., Shofer, F. S., Barg, F. K., & Meisel, Z. F. (2014). Evidence-based narratives to improve recall of opioid prescribing guidelines: A randomized experiment. *Academic Emergency Medicine*, *21*(3), 244–249. https://doi.org/10.1111/acem.12326

† ‡ Koopman, E. M. (2016). *Reading suffering: An empirical inquiry into empathic and reflective responses to literary narratives* [Doctoral dissertation, Erasmus University Rotterdam]. ERMeCC – Erasmus Research Centre for Media, Communication and Culture. http://hdl.handle.net/1765/93344

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

‡ Kozak, S., & Martin-Chang, S. (2019). Preservice teacher knowledge, print exposure, and planning for instruction. *Reading Research Quarterly*, *54*(3), 323–338. https://doi.org/10.1002/rrq.240

Lamarque, P., & Olsen, S. H. (1994). *Truth, fiction, and literature: A philosophical perspective*. Oxford University Press.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, *44*(2), 325–343. https://doi.org/10.3758/s13428-011-0146-0

‡ Lenhart, J., Dangel, J., & Richter, T. (2020). The relationship between lifetime book reading and empathy in adolescents: Examining transportability as a moderator. *Psychology of Aesthetics, Creativity, and the Arts*, *16*(4), 679–693. https://doi.org/10.1037/aca0000341

‡ Lenhart, J., Richter, T., Appel, M., & Mar, R. A. (2023). Adolescent leisure reading and its longitudinal association with prosocial behavior and social adjustment. *Scientific Reports*, *13*(1), 9695. https://doi.org/10.1038/s41598-023-35346-7

† ‡ 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*. https://doi.org/10.1037/aca0000533

‡ Lenhart, J., & Richter, T. (2023). *Does watching a limited amount of fictional TV series in a naturalistic setting increase social-cognitive skills?* Unpublished manuscript.

Liberman, M. (2013). *Annals of overgeneralization*. Language Log. https://languagelog.ldc.upenn.edu/nll/?p=7715

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis.* SAGE.

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

† Małecki, W., Pawłowski, B., Cieński, M., & Sorokowski, P. (2018). Can fiction make us kinder to other species? The impact of fiction on pro-animal attitudes and behavior. *Poetics, 66,* 54–63. https://doi.org/10.1016/j.poetic.2018.02.004

‡ Seven unpublished studies by Mar

Mar, R. A. (2018a). 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. (2018b). 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. (2007). *Stories and the simulation of social experience: Neuropsychological evidence and social ramifications* [Doctoral dissertation, University of Toronto]. Tspace Repository. https://hdl.handle.net/1807/119125

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. B., Paz, J. D., & 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

‡ Mar, R. A., & Rain, M. (2015). Narrative fiction and expository nonfiction differentially predict verbal ability. *Scientific Studies of Reading*, *19*(6), 419–433. https://doi.org/10.1080/10888438.2015.1069296

† Mara, M., & Appel, M. (2015). Science fiction reduces the eeriness of android robots: A field experiment. *Computers in Human Behavior*, *48*, 156–162. https://doi.org/10.1016/j.chb.2015.01.007

† Maria, K., & Junge, K. (1993). *A comparison of fifth graders' comprehension and retention of scientific information using a science textbook and an informational storybook* (ED364864). ERIC. https://files.eric.ed.gov/fulltext/ED364864.pdf

Marsh, E. J., & Fazio, L. K. (2006). Learning errors from fiction: Difficulties in reducing reliance on fictional stories. *Memory & Cognition*, *34*(5), 1140–1149. https://doi.org/10.3758/BF03193260

Marsh, E., Meade, M. L., & Roediger III, H. L. (2003). Learning facts from fiction. *Journal of Memory and Language*, *49*(4), 519–536. https://doi.org/10.1016/S0749-596X(03)00092-5

‡ Martin-Chang, S., Kozak, S., Levesque, K., Calarco, N., & Mar, R. A. (2021). What's your pleasure? Exploring the predictors of leisure reading for fiction and nonfiction. *Reading and Writing, 34*(6), 1387–1414. https://doi.org/10.1007/s11145-020-10112-7

† Maxim, I. R. (2022). *Effects of fiction reading on episodic memory* [Unpublished master's thesis]. Universitetet I Oslo.

† Mazor, K. M., & Billings-Gagliardi, S. (2003). Does reading about stroke increase stroke knowledge? *Patient Education and Counseling*. https://doi.org/10.1016/s0738-3991(02)00218-5

‡ Miall, D. S., & Kuiken, D. (1995). Aspects of literary response: A new questionnaire. *Research in the Teaching of English, 29(1),* 37-58. https://www.jstor.org/stable/40171422

Moll, J., Zahn, R., De Oliveira-Souza, R., Krueger, F., & Grafman, J. (2005). The neural basis of human moral cognition. *Nature Reviews Neuroscience*, *6*(10), 799–809. https://doi.org/10.1038/nrn1768

Moosbrugger, H., & Kelava, A. (Eds.) (2020). *Testtheorie und Fragebogenkonstruktion [Test theory and questionnaire construction]* (3th rev. ed.). Springer.

Mumford, M. D. (2003). Where have we been, where are we going? Taking stock in creativity research. *Creativity Research Journal*, *15*(2–3), 107–120. https://doi.org/10.1080/10400419.2003.9651403

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

Oakley, B. F. M., Brewer, R., Bird, G., & Catmur, C. (2016). Theory of mind is not theory of emotion: A cautionary note on the Reading the Mind in the Eyes Test. *Journal of Abnormal Psychology, 125*(6), 818–823. https://doi.org/10.1037/abn0000182

Oatley, K., Dunbar, R., & Budelmann, F. (2018). Imagining possible worlds. *Review of General Psychology*, *22*(2), 121–124. https://doi.org/10.1037/gpr0000149

† Orellana, L. (2019). *Fiction as mediated contact: Mechanisms of fiction associated with lower prejudice towards sexual and gender minorities* [Unpublished doctoral dissertation]. University of Sheffield.

† Orellana, L., Totterdell, P., & Iyer, A. (2020). The association between transgender-related fiction and transnegativity: Transportation and intergroup anxiety as mediators. *Psychology and Sexuality*, *13*(2), 228–239. https://doi.org/10.1080/19419899.2020.1759677

Østby, Y., Walhovd, K. B., Tamnes, C. K., Grydeland, H., Westlye, L. T., & Fjell, A. M. (2012). Mental time travel and default-mode network functional connectivity in the developing brain. *Proceedings of the National Academy of Sciences, 109*(42), 16800–16804. https://doi.org/10.1073/pnas.1210627109

† Panero, M. E., Weisberg, D. S., Black, J. M., 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

† Petterson, A., Currie, G., Friend, S., & Ferguson, H. J. (2022). The effect of narratives on attitudes toward animal welfare and pro-social behaviour on behalf of animals: Three pre-registered experiments. *Poetics*, *94*, 101709. https://doi.org/10.1016/j.poetic.2022.101709

‡ Pfost, M., Dörfler, T., & Artelt, C. (2013). Students' extracurricular reading behavior and the development of vocabulary and reading comprehension. *Learning and Individual Differences*, *26*, 89–102. https://doi.org/10.1016/j.lindif.2013.04.008

† ‡ Pigden, A. R. (2021). *Is the pen mightier than the sword? Testing the effects of fiction on empathy and aggression, and introducing a hybrid model of media effects* [Doctoral dissertation, University of Otago]. OUR Archive. http://hdl.handle.net/10523/12085

† Pino, M. C., & Mazza, M. (2016). The use of "Literary Fiction" to promote mentalizing ability. *PLOS ONE*, *11*(8), e0160254. https://doi.org/10.1371/journal.pone.0160254

Prentice, D. A., Gerrig, R. J., & Bailis, D. S. (1997). What readers bring to the processing of fictional texts. *Psychonomic Bulletin & Review*, *4*(3), 416–420. https://doi.org/10.3758/BF03210803

Preston, S. D., Ermler, M., Lei, Y., & Bickel, L. (2020). Understanding empathy and its disorders through a focus on the neural mechanism. *Cortex, 127*, 347–370. https://doi.org/10.1016/j.cortex.2020.03.001

Quinlan, J. A., Padgett, J. K., Khajehnassiri, A., & Mar, R. A. (2022). 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

Rapp, D. N., Hinze, S. R., Slaten, D. G., & Horton, W. A. (2014). Amazing stories: Acquiring and avoiding inaccurate information from fiction. *Discourse Processes*, *51*(1–2), 50–74. https://doi.org/10.1080/0163853x.2013.855048

Raven, I. C. (1962). *Advanced Progressive Matrices* (Set Il). H. K. Lewis

† ‡ Redman, K. (2017). *Fiction reading and sexism: Exploring the effects of fiction reading and transportation on sexist attitudes* [Doctoral dissertation, University of Otago]. OUR Archive. http://hdl.handle.net/10523/7411

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86*(3), 638–641. https://doi.org/10.1037/0033-2909.86.3.638

Rothstein H. R. (2008). Publication bias as a threat to the validity of meta-analytic results. *Journal of Experimental Criminology*, 4(1), 61–81. https://doi.org/10.1007/s11292-007-9046-9

‡ Samur, D., Luminet, O., & Koole, S. L. (2017). Alexithymia predicts lower reading frequency: The mediating roles of mentalising ability and reading attitude. *Poetics, 65*, 1–11. https://doi.org/10.1016/j.poetic.2017.09.002

† 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 & Emotion, 32*(1), 130–144. https://doi.org/10.1080/02699931.2017.1279591

† Savitri, S. I., & Nuha, F. K. (2021). The role of narrative format in increasing narrative transport and empathy for fiction and non-fiction readers. *Jurnal Psikologi, 20*(2), 152-162.

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

Schultz, A. (2020). Sitting still and reading: Rethinking the role of literary fiction in civics education. *Educational Theory, 70*(2), 187-198. https://doi.org/10.1111/edth.12415

Schurz, M., Radua, J., Tholen, M., Maliske, L., Margulies, D., Mars, R., 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*. 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

† Seddio, K. (2017). *Wearing the inside out: The effects of exogenous oxytocin, reading, and stress on the expression of empathy for victims of trauma* [Unpublished doctoral dissertation]. University of North Texas.

Shamay-Tsoory, S. G., & Aharon-Peretz, J. (2007). Dissociable prefrontal networks for cognitive and affective theory of mind: A lesion study. *Neuropsychologia*, *45*(13), 3054–3067. https://doi.org/10.1016/j.neuropsychologia.2007.05.021

Sincoff, J. B., & Sternberg, R. J. (1987). Two faces of verbal ability. *Intelligence*, *11*(4), 263–276. https://doi.org/10.1016/0160-2896(87)90010-9

‡ Spear-Swerling, L., Brucker, P. O., & Alfano, M. C. (2010). Relationships between sixth-graders' reading comprehension and two different measures of print exposure. *Reading and Writing*, *23*(1), 73–96. https://doi.org/10.1007/s11145-008-9152-8

‡ Spear-Swerling, L., wie, Y., Dostal, H. M., & Hernandez, B. (2020). The print exposure of teacher candidates in relation to their achievement and self-ratings of early reading experiences. *Reading and Writing, 33*, 2097–2119. https://doi.org/10.1007/s11145-020-10035-3

‡ Stanisic, B. (2019). *Fantasy versus reality: How video game and book genres associate with creative thinking* [Unpublished master's thesis]. Linnaeus University.

‡ Stanovich, K. E., & Cunningham, A. E. (1992). Studying the consequences of literacy within a literate society: The cognitive correlates of print exposure. *Memory & Cognition*, *20*(1), 51–68. https://doi.org/10.3758/bf03208254

‡ Stansfield, J., & Bunce, L. (2014). The relationship between empathy and reading fiction: Separate roles for cognitive and affective components. *Journal of European Psychology Students, 5(3),* 9–18. https://doi.org/10.5334/jeps.ca

† Stern, S. (2020). *Using popular media to change attitudes and bolster knowledge about Autism Spectrum Disorder* [Unpublished doctoral dissertation]. The University of Oklahoma.

‡ Stern, S. M., Robbins, B., Black, J. M., & Barnes, J. L. (2019). What you read and what you believe: Genre exposure and beliefs about relationships. *Psychology of Aesthetics, Creativity, and the Arts*, *13*(4), 450–461. https://doi.org/10.1037/aca0000189

Summers, K. (2013). Adult reading habits and preferences in relation to gender differences. *Reference & User Services Quarterly, 52*(3), 243-249. https://www.jstor.org/stable/10.2307/refuseserq.52.3.243

‡ Sutton, N. (2021). *The exposure to literary fiction makes individuals more empathetic, fact or fiction.* [Unpublished undergraduate thesis]. National College of Ireland.

‡ Tabullo, A. J., Chiófalo, M. F., & Wainselboim, A. J. (2022). Reading comprehension in undergraduates during the COVID-19 pandemic. Associations with executive function difficulties, reading habits and screen times. *Research Square*. https://doi.org/10.21203/rs.3.rs-1804651/v1

‡ Tabullo, A. J., Jiménez, V. A. N., & García, C. S. (2018). Associations between fiction reading, trait empathy and theory of mind ability. *International Journal of Psychology and Psychological Therapy*, *18*(3), 357-370.

‡ Tabullo, A. J., Pithod, M., & Moreno, C. B. (2020). Associations between reading, comprehension, print exposure, executive functions and academic achievement in Argentinean university students. *Revista Neuropsicología, Neuropsiquiatría y Neurociencias, 20 (2),* 15-48. https://repositorio.uca.edu.ar/handle/123456789/13511

†‡ Takahashi, Y., Himichi, T., Masuchi, A., Nakanishi, D., & Ohtsubo, Y. (2023). Is reading fiction associated with a higher mind-reading ability? Two conceptual replication studies in Japan. *PLOS ONE, 18*(6), e0287542. https://doi.org/10.1371/journal.pone.0287542

‡ 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

Tebbe, E. A., Moradi, B., & Ege, E. (2014). Revised and abbreviated forms of the Genderism and Transphobia Scale: Tools for assessing anti-trans* prejudice. *Journal of Counseling Psychology, 61*(4), 581–592. https://doi.org/10.1037/cou0000043

Tierney, R. J., & LaZansky, J. (1980). The rights and responsibilities of readers and writers: A contractual agreement. *Language Arts, 57*(6), 606-613.

Tsou, A., Treadwell, J., Erinoff, E., & Schoelles, K. M. (2020). Machine learning for screening prioritization in systematic reviews: Comparative performance of Abstrackr and EPPI-Reviewer. *Systematic Reviews, 9*(1). https://doi.org/10.1186/s13643-020-01324-7

‡ Turner, R., & Felisberti, F. M. (2018). Relationships between fiction media, genre, and empathic abilities. *Scientific Study of Literature, 8*(2), 261–292. https://doi.org/10.1075/ssol.19003.tur

† ‡ Turner, R., & Vallée-Tourangeau, F. (2020). Fiction effects on social cognition. *Scientific Study of Literature, 10*(1), 64–97. https://doi.org/10.1075/ssol.19008.tur

‡ Turner, R., & Vallée-Tourangeau, F. (2023). Challenges of measuring empathic accuracy: A mentalizing versus experience-sharing paradigm. *British Journal of Social Psychology.* https://doi.org/10.1111/bjso.12612

Vallacher, R. R., & Wegner, D. M. (1989). Levels of personal agency: Individual variation in action identification. *Journal of Personality and Social Psychology*, *57*(4), 660–671. https://doi.org/10.1037/0022-3514.57.4.660

VandenBos, G. R. (2015). *APA dictionary of psychology*. American Psychological Association.

‡ Van Der Kleij, S. W., Apperly, I. A., Shapiro, L. R., Ricketts, J., & Hughes, C. (2022). Reading fiction and reading minds in early adolescence: A longitudinal study. *Journal of Experimental Child Psychology*, *222*, 105476. https://doi.org/10.1016/j.jecp.2022.105476

‡ Van Kuijk, I., Verkoeijen, P. P. J. L., Dijkstra, K., & Zwaan, R. A. (2018). The effect of reading a short passage of literary fiction on Theory of Mind: A replication of Kidd and Castano (2013). *Collabra*, *4*(1). https://doi.org/10.1525/collabra.117

‡ Vermeiren, H., Vandendaele, A., & Brysbaert, M. (2022). Validated tests for language research with university students whose native language is English: Tests of vocabulary, general knowledge, author recognition, and reading comprehension. *Behavior Research Methods*. https://doi.org/10.3758/s13428-022-01856-x

† ‡ Vezzali, L., McKeown, S., McCauley, P., Stathi, S., Di Bernardo, G. A., Cadamuro, A., Cozzolino, V., & Trifiletti, E. (2021). May the odds be ever in your favor: The Hunger Games and the fight for a more equal society. (Negative) Media vicarious contact and collective action. *Journal of Applied Social Psychology*, *51*(2), 121–137. https://doi.org/10.1111/jasp.12721

† Vezzali, L., Stathi, S., & Giovannini, D. (2012). Indirect contact through book reading: Improving adolescents' attitudes and behavioral intentions toward immigrants. *Psychology in the Schools*, *49*(2), 148–162. https://doi.org/10.1002/pits.20621

‡ Vezzali, L., Stathi, S., Giovannini, D., Capozza, D., & Trifiletti, E. (2015). The greatest magic of Harry Potter: Reducing prejudice. *Journal of Applied Social Psychology, 45*(2), 105–121. https://doi.org/10.1111/jasp.12279

Viechtbauer, W. (2010). Conducting meta-analyses in *R* with the metafor package. *Journal of Statistical Software*, *36*(3). https://doi.org/10.18637/jss.v036.i03

Vrasidas, C., Avraamidou, L., Theodoridou, K., Themistokleous, S., & Panaou, P. (2015). Science Fiction in education: Case studies from classroom implementations. *Educational Media International*, *52*(3), 201–215. https://doi.org/10.1080/09523987.2015.1075102

† Walkington, Z., Wigman, S. A., & Bowles, D. S. (2020). The impact of narratives and transportation on empathic responding. *Poetics*, *80*, 101425. https://doi.org/10.1016/j.poetic.2019.101425

Wallace, B. C., Small, K., Brodley, C. E., Lau, J., & Trikalinos, T. A. (2012). Deploying an interactive machine learning system in an evidence-based practice center: Abstrackr. *Proceedings of th^e 2nd ACM SIGHIT International Health Informatics Symposium, January 2012*, 819–824. https://doi.org/10.1145/2110363.2110464

‡ Waytz, A., Hershfield, H. E., & Tamir, D. I. (2015). Mental simulation and meaning in life. *Journal of Personality and Social Psychology*, *108*(2), 336–355. https://doi.org/10.1037/a0038322

‡ West, R. G., Stanovich, K. E., & Mitchell, H. H. (1993). Reading in the real world and its correlates. *Reading Research Quarterly*, *28*(1), 34. https://doi.org/10.2307/747815

‡ Whalen, L. M. (2010). *Empathy and reading of narrative fiction among community college students* (Publication No 3397454) [Doctoral dissertation, Capella University]. ProQuest Dissertations.

Wheeler, C., Green, M. C., & Brock, T. C. (1999). Fictional narratives change beliefs: Replications of Prentice, Gerrig, and Bailis (1997) with mixed corroboration. *Psychonomic Bulletin & Review*, *6*(1), 136–141. https://doi.org/10.3758/BF03210821

White, S. J., Coniston, D., Rogers, R., & Frith, U. (2011). Developing the Frith-Happé animations: A quick and objective test of Theory of Mind for adults with autism. *Autism Research*, *4*(2), 149–154. https://doi.org/10.1002/aur.174

Williams, P. G., Rau, H. K., Suchy, Y., Thorgusen, S. R., & Smith, T. W. (2017). On the validity of self-report assessment of cognitive abilities: Attentional control scale associations with cognitive performance, emotional adjustment, and personality. *Psychological Assessment*, *29*(5), 519–530. https://doi.org/10.1037/pas0000361

† Willis, M. N. (1998). *Effects of reading narrative and expository informational trade books on third-grade students' science knowledge and motivation to extend science knowledge* (Publication No. 9828334) [Doctoral dissertation, University of Houston]. ProQuest Dissertations.

‡ Wimmer, L., Currie, G., Friend, S., & Ferguson, H. J. (2021a). Testing correlates of lifetime exposure to print fiction following a multi-method approach: Evidence from young and older readers. *Imagination, Cognition and Personality*, *41*(1), 54–86. https://doi.org/10.1177/0276236621996244

† Wimmer, L., Currie, G., Friend, S., & Ferguson, H. J. (2022a). The effects of reading narrative fiction on social and moral cognition. *Scientific Study of Literature*, *11*(2), 223–265. https://doi.org/10.1075/ssol.21010.wim

† ‡ Wimmer, L., Currie, G., Friend, S., & Ferguson, H. J. (2022b). Opening the closed mind? Effects of reading literary fiction on need for closure and creativity. *Creativity Research Journal*, 1–18. https://doi.org/10.1080/10400419.2022.2087309

‡ Wimmer, L., El-Salahi, L., Lee, W. H., & Ferguson, H. J. (2022c). Narrativity and literariness affect the aesthetic attitude in text reading. *Empirical Studies of the Arts*, *41*(1), 231–258. https://doi.org/10.1177/02762374221095482

‡ Wimmer, L., & Ferguson, H. J. (2022). Testing the validity of a self-report scale, author recognition test, and book counting as measures of lifetime exposure to print fiction. *Behavior Research Methods*, *55*(1), 103–134. https://doi.org/10.3758/s13428-021-01784-2

‡ Wimmer, L., Friend, S., Currie, G., & Ferguson, H. J. (2021b). Reading fictional narratives to improve social and moral cognition: The influence of narrative perspective, transportation, and identification. *Frontiers in Communication*, *5*. https://doi.org/10.3389/fcomm.2020.611935

† Wolfe, M. S., & Mienko, J. A. (2007). Learning and memory of factual content from narrative and expository text. *British Journal of Educational Psychology*, *77*(3), 541–564. https://doi.org/10.1348/000709906x143902

† Wolfe, M. S., & Woodwyk, J. M. (2010). Processing and memory of information presented in narrative or expository texts. *British Journal of Educational Psychology*, *80*(3), 341–362. https://doi.org/10.1348/000709910x485700

† Wulandini, I. A. Kuntoro, & Handayani, E. (2018). The effect of literary fiction on school-aged children's Theory of Mind (ToM). In A. A. Ariyanto, H. Muluk, P. Newcombe, F. P. Piercy, E. K. Poerwandari, & S. H. R. Suradijono (Eds.), *Diversity in unity: Perspectives from psychology and behavioral sciences* (pp. 159-166). Routledge.

‡ Xu, W., Gao, Q., Yu, X., Guo, J., & Wang, R. (2022). Cumulative reading engagement predicts individual sensitivity to moral judgment: The mediating role of social processing tendencies. *Studia Psychologica*, *64*(3), 268–282. https://doi.org/10.31577/sp.2022.03.853

Young, L., & Koenigs, M. (2007). Investigating emotion in moral cognition: A review of evidence from functional neuroimaging and neuropsychology. *British Medical Bulletin, 84*(1), 69–79. https://doi.org/10.1093/bmb/ldm031

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

Zurek, P. P., & Scheithauer, H. (2017). Towards a more precise conceptualization of empathy: An integrative review of literature on definitions, associated functions, and developmental trajectories. *International Journal of Developmental Science, 11*(3–4), 57–68. https://doi.org/10.3233/DEV-16224