The Benefits of Fiction-engagement for Empathic Abilities:

A Multidimensional Approach

Rose TURNER

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy (PhD) at the Department of Psychology, Kingston University London.

February 2020

Associated Publications

Parts of this dissertation have been reproduced in the following publications:

Turner, R., & Vallée-Tourangeau, F. (in press). Fiction effects on social cognition: Varying narrative engagement with cognitive load. *Scientific Study of Literature, 10.* Copyright John Benjamins Publishing Company.

Turner, R., & Felisberti, F. (2018). Relationships between fiction media, genre, and empathic abilities. *Scientific Study of Literature, 8,* 261-292. http://doi.org/10.1075/ssol.19003.tur. Copyright John Benjamins Publishing Company.

Turner, R. & Felisberti, F. M. (2017). Measuring mindreading: A review of behavioral approaches to measuring "theory of mind" in neurologically typical adults. *Frontiers in Psychology, 8,* 1-7. http://doi.org/10.3389/fpsyg.2017.00047. Licensed by Creative Commons (CC-BY: https://creativecommons.org/licenses/by/4.0/).

What a miracle it is that out of these small, flat, rigid squares of paper unfolds world after world after world, worlds that sing to you, comfort and quiet or excite you. Books help us understand who we are and how we are to behave. They show us what community and friendship mean; they show us how to live and die.

—Anne Lamott, *Bird By Bird,* 1994

Table of Contents

List of Tables

List of Figures

Acknowledgments

Abstract

The processes involved in empathy, the ability to appreciate others' inner experiences and respond appropriately to them, are central to the formation and maintenance of successful interpersonal relationships and communities (e.g., Castano, 2012). These skills typically emerge in childhood but can also be developed in adults (Teding van Berkhout & Malouff, 2016). Engagement with fiction may enhance adults' empathic skills because readers mentally simulate the social experiences depicted in stories (Oatley, 1999). Several studies have identified positive relationships between exposure to fiction and empathic abilities (Mumper & Gerrig, 2017), whereas causal findings are more mixed (see Dodell-Feder & Tamir, 2018), and this may reflect heterogeneity across both fiction stimuli and empathy measures.

The present research took a multidimensional approach to investigating the nature of relationships between fiction and empathic abilities. Study 1 examined correlations between self-report empathic abilities and fiction habits. Participants ($N = 404$) completed a multidimensional task measure of fiction media-exposure and answered questions about fiction-engagement and empathic tendencies. Results revealed divergent associations between narrative modes and empathic abilities, and fiction media-exposure positively predicted the tendencies to become absorbed in narratives and to behave altruistically. Study 2 ($N = 308$) assessed the relationship between fiction-exposure and performance on a behavioural measure of empathic accuracy (the ability to accurately interpret mental state content) when using mentalising or experience-sharing inferencing processes. Results showed that the two strategies entailed similar levels of error but in opposite directions. Empathic accuracy varied as a function of target and valence and was positively predicted by lifetime fiction-exposure. Study 3 investigated the causal impact of immersion. An initial pilot study, a text pretest, two manipulation pilots, and an experiment (total $N = 224$), were conducted. Ultimately, immersion levels, measured across three dimensions, were not successfully manipulated.

Immersion dimensions correlated with self-report and behavioural empathic ability measures, and an exploratory analysis revealed an effect of reading on empathic accuracy for story characters' mental states. Collectively, the studies provide support for the hypothesis that fiction-exposure and empathic abilities are associated, but limited evidence of causation. Methodological limitations, other influential variables, and research implications are discussed. The assumption that fiction and empathy are beneficial is critiqued, and future research avenues suggested.

Chapter 1: Introduction to the Thesis

Chapter 1 provides an overview of the dissertation. It introduces the research topic: the relationship between fiction and empathy, outlines the research rationale, and establishes the organisation of the dissertation. The importance of a psychological account of fiction-engagement is identified, supported by evidence from scholarship, as well as examples of social initiatives operating in the UK, USA and Canada that use fiction-based applications to enhance interpersonal skills. The empathy construct is defined as comprising three dissociable sub-components: mentalising, experience-sharing, and prosocial concern. Mentalising and experience-sharing are contrasted as strategies for "empathic accuracy": the accurate interpretation of others' inner states, which can lead to prosocial concern. Prosocial concern is established as the basis of prosocial behaviours. The case for studying empathic processes in neurologically typical (NT) adults is outlined, and approaches to measurement, including fiction-based tools, are discussed. A brief overview of empirical evidence showing associations between fiction and empathy is presented, and inconsistencies in the literature are discussed in terms of the multidimensionality of both fiction-engagement and the empathy construct. It is proposed that fiction formats differentially relate to dimensions of empathy, and that fiction-engagement may particularly support the experience-sharing strategy. The chapter concludes with an outline of the dissertation structure, including a summary of each chapter's contents.

1.1 Fiction and Society

> We do not have a user manual from which to write down an exhaustive set of values for a culture [but] we do have collective stories told by those belonging to different cultures. (Riedl & Harrison, 2016, p. 106)

Humans are the only animals that create, tell, buy and sell stories. The time spent engaging with the *Harry Potter* franchise alone is estimated at more than 235,000 years (Barnes, 2012). Fiction is big business (Nettle, 2005), and with technological developments

allowing people to become immersed in stories through portable devices, virtual reality, 3D cinema and an increasing multitude of television channels, its ubiquity seems boundless. Considering the time and money spent on fiction (Loughborough University, 2015; Office for National Statistics, 2012)—on books, radio, TV, apps, interactive media, and trips to the movies or theatre—it should offer some kind of benefit otherwise, through a process of cultural evolution, it would surely have ceased to exist (Nettle, 2005).

Several scholars have suggested that literature and the arts function to benefit society. Theatre practices in Ancient Greece promoted democracy via the staging of opposing viewpoints performed as dialogue, often supported by a chorus that emphasised prevalent moral views (Calame, 1999). The humanitarian reform of the 1700s, which included abolitionism, improvements in workers' conditions and healthcare, was preceded by the proliferation of literacy that followed the advent of mechanical printing (Pinker, 2011). Literature may develop social justice (e.g., Nussbaum, 1990, 1995), functioning as a "moral laboratory", wherein social attitudes can be explored and refined (Hakemulder, 2000), and it provides an opportunity to emotionally engage with social events without real-word consequences (Keen, 2007). By encouraging identification with others, literature, and the arts more broadly, can cultivate social intelligence (Oatley, 1999; Zunshine, 2006), and help to produce more cooperative, ethical citizens (Bazalgette, 2017).

The ethos that stories can develop social skills has been adopted by several social enterprises aiming to encourage positive relationships among people, to increase prosocial, and reduce antisocial, behaviours. *Roots of Empathy*, a programme initiated in Canada and now implemented across eleven countries, uses fiction-engagement as part of a schools-based curriculum aimed at fostering kindness, and UK organisation *EmpathyLab* uses literature as the key tool in its social-awareness interventions for 4-11-year-olds. In prisoners, literacy is generally lower than in the rest of the population (e.g., Creese, 2015), and several

programmes aimed at offender-rehabilitation involve fiction: *Book Clubs for Inmates* in Canada, and *Prison Reading Groups* in the UK both use literature to develop participants' prosocial skills. *Prison Performing Arts* in the US and *Clean Break* in the UK employ theatre programmes aimed at supporting prisoners to develop the communication and collaboration skills required for societal reintegration and employment, and *The Reader* uses shared reading groups to reduce social isolation in UK criminal justice and health settings. *Ladder to the Moon's* programmes, endorsed by the *Care Quality Commission*, the independent regulator of health and adult social care services in England, use theatre and filmmaking to build relationships among staff and care home residents, and to facilitate the provision of person-centred care (Ladder to the Moon, 2015).

Despite anecdotal, academic and economic accounts of their positive impact (e.g., EmpathyLab, 2019; Hartley & Turvey, 2013; Johnson, Keen & Pritchard, 2011; Prison Reading Groups, 2016; Schonert-Reichl, Smith, Zaidman-Zait & Hertzman, 2012), the implementation of arts-based social initiatives remains sporadic and largely underfunded, and arts education is diminishing. Arts organizations in the USA and UK are increasingly reliant on private donors (MTM London, 2016; Woronkowicz, Nichols & Iyengar, 2012); in the UK, local authority spending on arts and culture declined by 17% from 2010 to 2015 (Harvey, 2016) and funding has continued to dwindle in the UK and throughout Europe (Inkei, 2019). The UK National Curriculum has marginalised the arts and humanities in favour of science, technology, engineering and maths (STEM) subjects (Department for Business, Innovation and Skills, 2015), and the English Baccalaureate, which the government plans to implement for all GCSE students by 2020, contains no arts subjects at all (Long & Bolton, 2017). Similarly, the US Common Core, which determines academic standards for mathematics and English language arts, has increased its emphasis on nonfiction, and this has led teachers to replace fiction with factual texts (Loveless, 2015). Articulation of the potential value of

literature and the arts seems to be missing something. The research reported in this dissertation aimed to provide some remedy, by contributing to a psychological account of the benefits of engagement with fictional narratives.

1.2 Understanding Other Minds

Fiction is inherently social; it centres on humans or humanlike agents, and their interactions with others (Mar & Oatley, 2008). While expository nonfiction, too, can feature humanlike agents, fiction is distinguished by its complex characterization and narrative structure, which may recruit, and consequently strengthen, the psychological mechanisms concerned with making sense of the social world.

The processes involved in navigating social phenomena are supported by empathy, a multifaceted construct which enables perceivers to interpret and respond to the mental states of other individuals (Zaki & Ochsner, 2012). The psychological literature lacks a standard definition of empathy; however rational, theory-based inferencing is often contrasted with more embodied, experiential processes (e.g., Coll, Viding, Rütgen, Silani, Lamm, Catmur & Bird, 2017; Goldman, 2006). A useful framework from the field of social neuroscience encompasses both theory-based and experiential facets of empathy and a route to prosocial behaviour (Zaki & Ochsner, 2012). It organises empathic processes into three components: (i) experience-sharing: vicariously experiencing a target's internal state, (ii) mentalising: explicitly interpreting a target's internal state (e.g., through reading facial cues), and (iii) prosocial concern: expressing motivation to improve a target's circumstances. According to this model, mentalising and experience-sharing represent two routes to the goal of interpreting others' inner states or "empathic accuracy" (Ickes, 1997). This, in turn, can activate prosocial motivation to help or to alleviate suffering.

Experimental studies of empathic accuracy have generally focused on children and groups with characteristic deficits, such as those with schizophrenia, psychopathic traits and

Autism Spectrum Disorder (ASD). Fewer studies have investigated empathic accuracy in neurologically typical (NT) adults. In response to a legitimate need for clinical and developmental perspectives, this agenda has facilitated the identification of diagnostic markers, but it also reflects a general assumption that core empathic skills are acquired in childhood and cannot be improved upon in NT adults. However, empathic accuracy varies between NT adults, can change throughout the lifespan (Duval, Piolino, Bejanin, Eustache & Desgranges, 2010; Happé, Winner & Brownell, 1998; Maylor, Moulson, Muncer, & Taylor, 2002), and is implicated in interpersonal relationships and prosocial behaviour (Castano, 2012; Paal & Bereczkei, 2007). Examining how it may be enhanced represents an important area of enquiry.

Due to the emphasis on development and disorders, most well-established tests of empathic accuracy show ceiling effects with NT adults. Consequently, studies examining nuanced variation at the upper echelons of the ability have necessitated more complex narrative-based tasks, using written prose (e.g., Dodell-Feder, Lincoln, Coulson & Hooker, 2013), film (e.g., Dziobek et al., 2006) and virtual reality stimuli (e.g., Spiers & Maguire, 2006). Sensitive, fiction-based measures enable the testing of participants' abilities to interpret naturalistic social scenarios. Providing physical, verbal and contextual information, such stimuli are more akin to real-world environments than some of the more established unidimensional tools (such as facial emotion recognition tests). Participants' accuracy in attributing thoughts and feelings to narrative characters is considered to be indicative of their general empathic skills. These approaches, therefore, hinge on the assumption that the capacity to comprehend the experiences of fictional characters corresponds to the ability to comprehend others in the real world.

1.3 Fiction as Simulation of the Social World

The view that fiction and reality are comparable in terms of social stimuli represents a key tenet of research examining fiction effects on interpersonal skills. Mar and Oatley have proposed that fiction constitutes a simulation of the social world (Mar & Oatley, 2008; Oatley, 1999, 2011b; Oatley & Djikic, 2017): a surgeon might simulate a complicated procedure before operating on a living person, or a research student might simulate a dataset before commencing data collection and analysis. In the same vein, fiction enables readers to imaginatively experience social interactions without directly participating in them. Mar and Oatley argued that tracking the experiences, intentions and emotions of fictional characters recruits and cultivates readers' empathic faculties, and that its capacity to develop these skills represents the core function of fiction.

This claim has received support from three areas of research: neuroscience has shown that similar brain regions are active when reading about an action, compared to when performing that same action (e.g., Speer, Reynolds, Swallow & Zacks, 2009), correlational research has revealed that, compared to nonfiction-readers, fiction-readers tend to demonstrate stronger empathic abilities (Mumper & Gerrig, 2017), and causal studies have indicated that reading about fictional characters can enhance mental state understanding (Dodell-Feder & Tamir, 2018).

Kidd and Castano (2013) conducted five experiments in which the effects of literary fiction-reading (fiction by award-winning or canonical authors) were compared to the effects of nonfiction, popular fiction (Amazon bestsellers or stories from a popular fiction anthology), or no reading. Their study was widely publicised for its finding that reading "literary" fiction temporarily improved empathic accuracy, whereas popular fiction and expository nonfiction did not. Several news outlets picked up the story, leading to headlines such as: "Reading Fiction Makes You A Nicer Person" (Barras, 2013), "Now We Have Proof

Reading Literary Fiction Makes You A Better Person" (Schonfeld, 2013), and "For Better Social Skills Scientists Recommend A Little Chekhov" (Belluck, 2013). The authors did not claim that general fiction-reading improves social skills, nor that literary fiction can make someone "a better person", but they did argue that literary prose incorporates complex characters, which recruits the mentalising faculty, whereas popular fiction merely entertains.

Subsequent replication efforts failed to reproduce these findings (e.g., Camerer et al., 2018; Panero, Weisberg, Black, Goldstein, Barnes, Brownell & Winner, 2016; Samur, Tops & Koole, 2018; cf. Kidd & Castano's, 2018a, response). This cast doubt on the hypothesis that any single fiction-reading session immediately enhances people's capacity to understand real-world others. However, a positive correlation between lifetime exposure to fiction and empathy task performance was observed in the original and replication studies. Perhaps, then, fiction-engagement does not strengthen empathy, but empathic people tend to read more fiction. Alternatively, effects may be causal but not immediate, or the specific empathic components impacted by fiction may not have been probed by the study design (e.g., Pino & Mazza, 2016, showed some support for the causal hypothesis using alternative stimulus texts and measures).

1.4 Multidimensionality in Fiction and Empathy Research

Clarifying the association between fiction and empathy requires the operationalisation of both and raises two consequential issues: (i) the multidimensional nature of fiction-engagement, and (ii) the heterogeneity of the empathy construct. While most research has tested hypotheses concerning the value of reading, fiction is engaged with via a range of media channels. Evidence that exposure to nonfiction books negatively predicts empathic abilities (Mar, Oatley, Hirsh, dela Paz & Peterson, 2006) or, at least, results in a smaller positive effect (Mumper & Gerrig, 2017), indicates that positive relationships between fiction and empathy are not contingent on reading processes. If effects on empathic abilities are

related to the simulation of social content rather than to general reading processes, they should be observable across other media too.

Empathic abilities have been linked to engagement with stories via television (Black & Barnes, 2015a), movies (Mar, Tackett & Moore, 2010) and play-acting (Goldstein, Wu & Winner, 2009; Nettle, 2006). This raises the question, is the presentation channel (e.g., screen, paper, audio, live, interactive) simply a vehicle through which fictional stories impact people, or can the medium itself facilitate, alter or inhibit fiction effects? It could be that movies train viewers' abilities to decode facial expressions, whereas reading and play-acting bolster people's experience-sharing capacity by inviting them to take first-person perspectives on characters' inner experiences. In view of research indicating the effects of different stories and channels of engagement on empathy, investigating how far relationships between fiction formats and empathic abilities differ could prove informative for both science and social enterprise.

The content that influences empathy does not appear to be limited to acclaimed literary prose: research has shown that popular fiction is also associated with empathy (Fong, Mullin & Mar, 2013) and moral reasoning (Black, Capps & Barnes, 2018), and that certain social themes can motivate prosocial behaviour (e.g., Koopman, 2015). Fiction is multidimensional in terms of presentation format as well as story content. By examining both, researchers could extend current knowledge of the antecedents and consequences of fiction-engagement processes.

If fiction effects on empathy are not uniquely accounted for by literary devices, other aspects of engagement must be at play. The narrative persuasion literature provides a possible solution to this problem. Engaging with narratives can alter people's attitudes and beliefs, and one aspect of engagement that has been shown to augment this effect is "transportation" (Green & Brock, 2000) or the experience of becoming immersed in a story. If this experiential

process can moderate the relationship between fictional content and real-world beliefs, it may also impact fiction effects on empathic abilities. This idea is supported by some evidence that low levels of immersion in a story lead to lower scores on empathy measures (Bal & Veltkamp, 2013).

The relationship between fiction and empathy, therefore, may not centre on an elaborative, analytical approach to complex and contradictory characters, but rather on the process of becoming deeply immersed in storyworlds. Immersion is another multifaceted construct. People may feel emotionally moved by a story, produce vivid mental images of its environment, or identify with its characters, and these dimensions may differently affect fiction-engagement outcomes. A granular approach to measuring fiction, and examining levels of immersion, would contribute to knowledge of the process through which fiction may enhance empathy.

The next major consideration for fiction and empathy research is the selection of measures used to test empathic acumen. Studies examining associations with fiction have included self-report and behavioural tests of emotion recognition (e.g., Pino & Mazza, 2016; Kidd & Castano, 2013), perspective-taking (Mar et al., 2006), experience-sharing and concern (e.g., Bal & Veltkamp, 2013; Mar et al., 2006). The range of approaches is noteworthy, because processes such as decoding facial expressions, perspective-taking, belief attribution and emotion-sharing can dissociate; proficiency in one area does not amount to proficiency across all. Not only do NT adults show significant variation across empathic domains (Cox, Uddin, Di Martino, Castellanos, Milham & Kelly, 2012), selective deficits have been well-documented in clinical and developmental populations. For example, alexithymia is typified by specific difficulties in recognising and describing emotions, people high in psychopathic traits may accurately attribute thoughts and feelings while lacking the emotion-sharing capacity, and people living with dementia can show impaired perspective-taking and belief

attribution abilities alongside relatively intact emotion recognition (Freedman, Binns, Black, Murphy & Stuss, 2013).

The lack of correlation among empathy measures was observed by Mar et al. (2006) in their study of associations between fiction-exposure and empathy. They found that the ability to ascribe mental state terms to actors' faces and the ability to interpret a social scene performed by actors were not associated. The authors suggested that these two performance-based measures of social ability may have recruited contrasting empathic strategies: the first task involved ascribing mental state terms to static photographs, and may entail the mentalising component of empathy, whereas the video task required decoding of dynamic cues, and may engage a more embodied, experience-sharing approach.

1.5 Two Strategies for Empathic Accuracy

Children around the world are familiar with the maxim: "treat others as you would like to be treated yourself", and empathy is also appealed to in order to shape adult behaviour ("please leave this bathroom as you would like to find it"). People are enjoined to empathise with the experiences and preferences of others by using their own as reference points. It is perhaps unsurprising, then, that the extent to which people empathise with others can depend on how far they identify with them (e.g., Davies, 1993).

Perceived similarity to another person not only impacts the degree of empathy, but also *how* people empathise: whether they tend to adopt experience-sharing or mentalising strategies (dissimilar others are less likely to invite experience-sharing; Zaki & Ochsner, 2012). Thus, different individuals' empathic strategies may vary in response to the same target. This is noteworthy because research has shown that mentalising and experience-sharing strategies for empathic accuracy are not equal. Zhou, Majka and Epley (2017) asked participants to guess people's emotional responses to a range of positive, negative and neutral images. Participants were assigned to one of three conditions, (i) simulation (experience-

sharing), (ii) theorisation (mentalising), and (iii) simultaneous (control). In the simultaneous condition, participants viewed a split screen showing the images that the targets viewed, alongside synchronised video footage of their facial reactions. In the theorisation condition, participants saw only the facial reaction video footage, and in the simulation condition, they saw only the images. Using experience-sharing, participants were significantly more accurate. However, when offered a financial incentive (extra cash for reaching the 80th percentile), and given the choice between simulation and theorisation, participants tended to select the less successful theorisation condition. The authors suggested that people tend to underestimate their ability to share others' experiences, because they overestimate *dissimilarity* between self and others.

According to Singer (1981/2011), by default, people empathise with only a small circle or "ingroup" of family and friends. Although ingroups have expanded as society has become more global to encompass other geographical, gender and ethnic groups, people tend to show bias when interpreting the experiences of "outgroup" members (e.g., Riva & Andrighetto, 2012). Sharing in the experiences of other agents—including fictitious others—appears to facilitate expansion of the "empathy circle" (Pinker, 2011; Singer, 1981/2011), and stories have been shown to increase empathy for marginalised individuals (Batson et al., 1997). The process through which this occurs remains opaque; however, if fiction-engagement enables readers to simulate the experiences of distant others, it may particularly strengthen the experience-sharing empathic strategy. Perhaps, then, frequent fiction-readers would be more likely to select into Zhou et al.'s (2017) simulation condition, and to perform better within it.

1.6 Fiction, Experience-sharing and Real-world Others

Learning to share in the experiences of fictional characters may underpin the benefits of fiction-based social initiatives. In successful prison reading groups, offenders develop an appreciation for the experiences of story characters, which strengthens their understanding of

themselves and their communities (Prison Reading Groups, 2016). In the care sector, the person-centred approach advocated by organisations such as *Ladder to the Moon* requires, by definition, the ability to interpret the unique experiences and preferences of individuals (e.g., Health Foundation, 2016). Following an intervention in which staff and residents co-produced a creative film project, a care home manager described improvements in relationships among staff members: "after the programme we understood each others' roles more and supported each other. I think we've been more reflective about the feelings of others" (Lyons & Gage, 2014, p. 3). After participating in *Empathy Lab's* Empathy Day, pupil "Romy" explained that: "if a character is lonely in a book, it will make you think about what it feels like to be lonely". "Gracie" added:

> When I'd finished the book, I thought a lot about it. One day when I was in the playground I saw a girl who looked really lonely, she was sitting down on her own, so I went and played with her. The book caused me to change my behaviour. (EmpathyLab, 2017)

Fiction-engagement appears to represent a method for activating empathic understanding and generating prosocial motivation. However, questions remain about the forms of fiction that benefit empathic abilities, and the empathic dimensions and strategies that can be cultivated. By addressing these questions, the present research aims to contribute to a granular understanding of the relationships between fictional stories and components of empathy.

1.7 Summary

Mentalising and experience-sharing represent alternative empathic processes that can be recruited for the same purpose: to accurately interpret the inner states of other individuals. In turn, these processes can motivate prosocial concern and behaviour, and so developing an understanding of how empathy may be enhanced is important for NT adults, as well as for clinical and developmental groups. The claim that fiction enables readers to simulate social experiences, which may help to develop their interpersonal skills, has received support from

empirical psychology and anecdotal evidence from real-world social enterprise. Familiarity with fiction positively predicts both self-report and performance-based empathic accuracy, and several social initiatives have reported positive outcomes for participants enrolled in fiction-based projects aimed at developing social understanding and prosociality.

Experimental studies examining the immediate, causal effects of fiction-engagement on empathy have yielded mixed findings. This may reflect the multidimensionality of the empathy construct and heterogeneity across fiction presentations. Empathy comprises dissociable components and processes, and fiction encompasses a range of narrative modes, which may entail different levels of immersion. Causal studies have tended to assess explicit mentalising; however, if fiction enables people to simulate the experiences of dissimilar others, it may also help to expand people's real-life empathic engagement beyond similar, ingroup members.

1.8 Organisation of the Dissertation

This dissertation addresses the general question: *How does engaging with fiction benefit empathic abilities in NT adults?* In Chapters 2-4, the relevant constructs are defined, and the research rationale and objectives are established. Chapter 2 defines and examines the empathy construct, presents a case for studying the phenomenon in NT adults, and reviews traditional approaches to measurement. Chapter 3 establishes the definition of fiction and presents a review of studies that have investigated relationships between fiction-engagement and empathic abilities. The research rationale, objectives and implications are summarised in Chapter 4, and the research hypotheses are articulated. Chapters 5-7 present the methods, results and analyses of the empirical work undertaken. Chapter 5 reports the first phase of the research: an exploration of relationships between fiction media and genre formats and self-report measures of empathy. Chapter 6 presents the second phase of the research, which examined the relationship between fiction-exposure, and mentalising versus experience-

sharing empathic strategies. Chapter 7 presents the third phase, which investigated the causal effects of immersion. In Chapter 8, the research findings are summarised, synthesised and discussed. Methodological limitations and potential contributions of the research are identified and avenues for further study are suggested.

Chapter 2: Empathy—A Multidimensional Construct

The willingness to understand others and respond prosocially to their needs is central to the formation of successful interpersonal relationships and cohesive communities (Castano, 2012; Paal & Bereczkei, 2007). Research has indicated that helpful, prosocial behaviours involve both empathic and altruistic personality dimensions (Penner, Fritzsche, Craiger & Freifeld, 1995), and that empathy can lead to altruistic behaviours (the willingness to help others with no expectation of reward; Penner et al., 1995) by enabling perceivers to infer the mental states of others (Batson, Duncan, Ackerman, Buckley & Birch, 1981). These skills typically develop in childhood (e.g., Perner & Wimmer, 1985) and so studies have tended to focus on children or adult groups with characteristic deficits. However, empathic abilities vary between neurologically typical (NT) adults and continue to develop across the lifespan (Duval et al., 2010; Happé et al., 1998; Maylor et al., 2002), and a growing body of literature has begun to address ways that adults' empathic skills may be enhanced. This chapter will define the empathy construct and model the relationship between empathic inferencing and prosocial behaviour. It will outline the case for studying empathic abilities in NT adults, and identify key paradigms aimed at measuring these skills.

2.1 The Empathy Concept

I do not ask the wounded person how he feels; I myself become the wounded person. (Whitman, *Leaves of Grass,* 1855/2017)

The term "empathy" has two etymological roots: the Greek "pathos", which refers to a quality that evokes sadness or pity, and the German "Einfühlung", literally translated as "in-feeling" (Waite, 2012). The latter term emerged in the 19th century from the work of the German Romantic philosophers. Robert Vischer (1873) first used the term "Einfühlung" in his doctoral thesis, *On the optical sense of form: A contribution to aesthetics*, where he used it to refer to the capacity to "feel into" objects: viewing an object such as a painting or sculpture

could produce sensations of pleasure which were then projected into and experienced within the object. The term was adopted by Theodor Lipps during the 1890s, when he developed and refined the concept as an aspect of perception. For Lipps, the self could be projected into an object or person in order to experience the world from "within" that other. He described watching a tightrope walker: "I am high up. I am transported there. Not next to the acrobat but exactly within him, where he is. This is the full meaning of *'Einfühlung'*" (Lipps, translated in Lanzoni, 2018, p. 33). Meanwhile, "sympathy" was becoming a central theme in English aesthetics. In *A treatise of human nature*, David Hume (Hume 1739-40/1985) proposed that "sympathy", the mechanism involved in sharing in the passions or views of others, contributed to all judgments of beauty. At the turn of the twentieth century, Edward Titchener, an English psychologist who studied under Wilhelm Wundt, coined the term "empathy". Whereas sympathy was to *feel with* or identify with another, "empathy" meant to *feel into* a person's experience or situation: "not only do I see gravity and modesty and pride and courtesy and stateliness, but I feel or act them in my mind's muscles" (Titchener, 1909, p. 21). By 1913, most psychologists had accepted Titchener's translation (although others had proposed "empathy" around this time, the term tends to be attributed solely to Titchener; Lanzoni, 2018).

For the past 100 years, empathy has been studied throughout social, personality and developmental psychology, and neuroscience, as well as philosophy and theology. Here, conceptualisations of empathy have evolved since its origins in aesthetics. The evocation of sadness implied via the etymological element "pathos" is no longer a prerequisite (although some conceptions of sympathy or concern assume a negative state), and projection into the experience of the target—the Einfühlung dimension—represents one theoretical standpoint on the ways that people empathise. In its broadest terms, empathy can be defined as "the notion of responsivity to the experience of another" (Davis, 1980, p. 3). Considering the range of

disciplines in which the construct has been operationalised and measured, it is unsurprising that a consistent definition is currently lacking (Coplan, 2004; Cuff, Brown, Taylor & Howat, 2016; Smith, 2017; see also Schaafsma, Pfaff, Spunt & Adolphs, 2015). De Vignemont and Singer (2006) commented that, "there are probably nearly as many definitions of empathy as people working on the topic" (p. 435).

In a review of the empathy concept, Cuff et al. (2016) identified forty-three discrete definitions, which differed along several dimensions: process (e.g., emotion contagion, sympathy), levels of distinction between self and other, congruence (emotions congruent with a target versus incongruent emotions), automaticity (automatic versus controlled processes), trait versus state abilities, and behavioural outcomes. Cuff et al. pointed out that most definitions assume the presence of an observed other, but that empathy can also be evoked by fictional or imaginary persons. They separated affective empathy (the experience of emotion in response to a stimulus) from cognitive empathy (the ability to understand another's feelings). Integrating the range of research definitions, Cuff et al. concluded that:

> There are functional differences between empathy and related concepts; empathy includes both cognitive and affective elements; the emotions of the target and observer are similar but not identical; other stimuli, such as imagination, can evoke empathy; a self/other distinction is maintained in empathy, although a degree of merging is necessary; empathy is affected by both trait and state influences; behavioural outcomes are not part of empathy itself; and finally, empathy is automatically elicited but is also subject to top-down controlled processes (Cuff et al., 2016, p. 7).[1]

---

[1] Cuff et al. (2016) proposed an affective definition (emotional response), generated via perception (direct or imagined) and understanding (which they referred to as "cognitive empathy"): they described an emotional response "dependent upon the interaction between trait capacities and state influences. Empathic processes are automatically elicited but are also shaped by top-down control processes. The resulting emotion is similar to one's perception (directly experienced or imagined) and understanding (cognitive empathy) of the stimulus emotion, with recognition that the source of the emotion is not one's own" (p. 7). In this way, their definition includes both cognitive and affective processes.

Cuff et al.'s (2016) review indicates that empathy is an important concept across several disciplines, but that agreement about precisely what it is, is lacking. The term has come to represent a broad range of processes, and these processes have also been clustered using other terms. "Theory of mind" (ToM; coined by Premack & Woodruff, 1978), has been widely used to refer to the capacity to recognise that others have different mental states to one's own, as well as the ability to interpret what they are (e.g., Baron-Cohen, Leslie & Frith, 1985; Wimmer & Perner, 1983). Apperly (2011) rejected this "linguistically awkward" term (p. 3), instead adopting "mindreading". He argued that "ToM" was suggestive of something that one *has* rather than something that one *does* and pointed out that whether or not a person has (or does not have) a theory about how the mind works represents only one theoretical possibility, whereas "mindreading" is theoretically neutral. "Mindreading" implies more agency than terms such as "mind perception" (e.g., Gray, Jenkins, Heberlein & Wegner, 2010); however, neither "ToM" nor "mindreading" capture the importance of accuracy. If it is accepted that the ability is dimensional rather than categorical—it is not something one has or does not have, but something one is able to do to different degrees—then approaches to measurement must index levels of accuracy.

Ickes (1997) distinguished between "empathic inference", the process of making sense of others' thoughts and feelings, and "empathic accuracy", the measure of one's skill in empathic inferencing. In this way, Ickes differentiated outcome from process; while a person can "theorise about", "mindread" or "perceive a mind" incorrectly, they cannot be empathically accurate incorrectly. Furthermore, his terminology invokes the aesthetic origins of the construct, where it is implicit that empathy is not necessarily constrained to present, real-world targets, but can also be entailed when imagining others, or engaging with fictional characters (as established in Cuff et al.'s, 2016, review). Therefore, the term "empathic accuracy" will be used here to refer to the measurable ability to accurately infer the inner

states of (real, fictional or imaginary) others, and "empathic inference" to refer to the process

or processes involved in interpreting those states. "Empathy" represents the umbrella term

(e.g., Pinotti & Salgaro, 2019) encompassing these dimensions and their subcomponents.

2.2 Components of Empathy

The extent that imaginative processes versus the process of instantiating another's inner

state are required for empathy has been the subject of much debate (for an overview, see

Gallagher & Gallagher, 2019). Zaki and Ochsner (2012) drew on findings from neuroscience

to outline a model of empathy that incorporated both sets of processes. They identified two

paths to empathic accuracy: mentalising (explicitly interpreting verbal and nonverbal cues),

and experience-sharing (vicariously sharing in the target's experience), each comprising a

range of sub-processes. The first domain, mentalising, includes theory of mind, perspective-

taking and cognitive empathy. Imagine the host of a party receiving a gift. One can work out

whether he or she liked the gift by interpreting the host's verbal and non-verbal cues and by

using prior knowledge about the recipient specifically, and gift-receiving reactions in general.

An alternative path, experience-sharing (the second domain), refers to the tendency to engage

the same neural systems when observing a state as when experiencing it first-hand. This path

enables the perceiver to draw on their own feelings, either about the gift itself or the host's

reaction, as the basis for accurately interpreting the host's experience. The distinction between

mentalising and experience-sharing is a question of *what* the host feels, versus *how* the host

feels it: experience-sharing requires not only a functional understanding of the mental state of

the target, but a matching of that state (see Smith, 2017).

Both mentalising and experience-sharing can lead to prosocial concern, the third facet

of Zaki and Ochsner's (2012) model, which is associated with prosocial behaviour. Here, the

party host may be disappointed by the gift, and one might wish to alleviate their negative

affect by providing a distraction, offering support or replacing the gift with something more

suitable.2 Via this route, either mentalising or experience-sharing could ultimately result in prosocial behaviour. This is in line with the empathy-altruism hypothesis (Batson, 1987; Batson, 2011; Batson et al., 1981; for an evaluation of research see Batson, Lishner & Stocks, 2015), which suggests that feeling concern for another can initiate an altruistic response.

Zaki and Ochsner's (2012) model has received support from neuroscientific research where mentalising and experience-sharing processes have been shown to initiate prosocial behaviour. However, the extent to which these processes are engaged, and their effects, depend on the context: for example, when people respond to explicit questions about targets' internal states, activity in brain areas associated with mentalising predicts helping (Harbaugh, Mayr & Burghart, 2007), whereas when watching a target in pain, activity in areas associated with experiencing that pain predicts helping (Hein, Silani, Preuschoff, Batson & Singer, 2010). This indicates that mentalising and experience-sharing processes represent two separable routes to empathic accuracy which, in turn, can initiate prosocial concern and behaviour.

For the purpose of this research, Zaki and Ochsner's (2012) model has been adapted to establish the two routes (mentalising versus experience-sharing) to the measurable empathic accuracy component, and to incorporate prosocial behavioural outcomes (Figure 1). As cognitive and affective components activate different neural processes (Singer, 2006), the model incorporates the cognitive and affective distinction present in Cuff et al.'s (2016) review, which helps to organise different conceptualisations of empathy (Zickfeld, Schubert, Siebt & Fiske, 2017). Whereas Zaki and Ochsner categorised affective empathy as an aspect of experience-sharing, but not a mentalising process, here cognitive and affective empathy are distinguished in terms of content (cognitive versus affective) and process (the target content

---

2 Some researchers have conceptualised the motivational component as a result of concern (or sympathy) but not as part of it (e.g., Batson et al., 1981). Others are in agreement with the idea that motivational concern is preceded by a cognitive (mentalising) component (e.g., Baron-Cohen & Wheelwright, 2004).

may be affective, and yet the process of making an inference about that content may be unemotional [mentalising]; whereas experience-sharing implies a match between the empathiser and target's emotional states). This provides a framework for examining mentalising versus experience-sharing abilities and their possible prosocial outcomes.
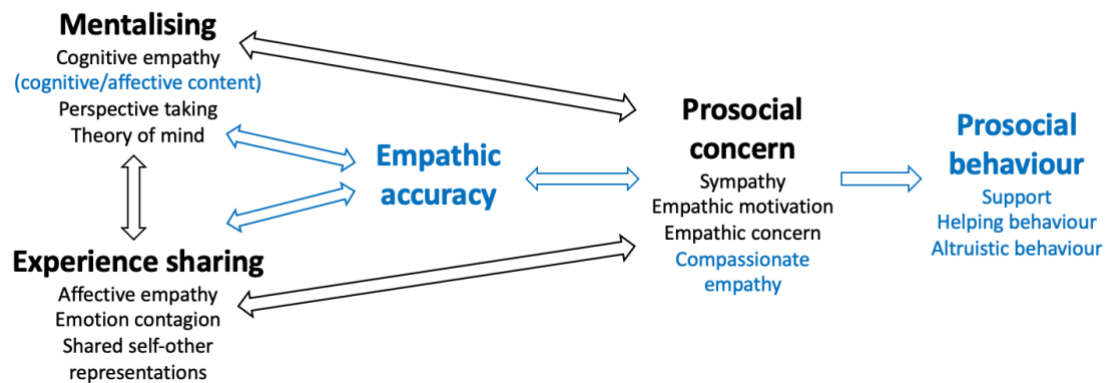
Figure 1. The major facets of empathy and their associated processes and terms adapted from "The Neuroscience of Empathy: Progress, Pitfalls and Promise", by J. Zaki and K. Ochsner, 2012, *Nature Neuroscience, 15*(5), p. 676, Nature Publishing Group (http://doi.org/10.1038/nn.3085). Neuroscience has focused on two processes (mentalising and experience-sharing), which can lead to prosocial concern. The model has been modified to incorporate cognitive and affective mentalising content, the measurable empathic accuracy component, and prosocial behaviour outcomes arising from the activation of prosocial concern. Additions to Zaki and Ochsner's model appear in blue.

The next section of this chapter is organised according to this framework. First, it sets out the case for studying empathic abilities in NT adults. Then it describes traditional paradigms for measuring abilities within the four domains: empathic accuracy via mentalising, empathic accuracy via experience-sharing, prosocial concern, and prosocial behaviour. Rather than present an exhaustive review of the measures available, this section will provide examples of the range of commonly used toolkits, including behavioural and self-report measures, from traditional vignette tasks to multidimensional approaches, that aim to index a range of processes via one task. This will lay the groundwork for interpreting the findings of research that employs these and similar measures in Chapter 3.

2.3 Empathic Abilities in Neurologically Typical Adults

Since Premack and Woodruff (1978) posed the question, "does the chimpanzee have a theory of mind?", research examining empathic accuracy has focused on child development, ASD and, more recently, clinical groups, whereas studies testing NT adults have been relatively infrequent. There are compelling arguments for investigating adults' empathic accuracy: the ability to accurately infer others' mental states changes across the lifespan (Duval et al., 2010; Happé et al., 1998; Maylor et al., 2002) and is positively associated with effective interpersonal relationships (Castano, 2012) and prosocial behaviour (Johnson, 2012; Paal & Bereczkei, 2007). Adults show different patterns of empathic ability compared to children. For example, in children, ASD and clinical samples, implicit empathic abilities, which are measured indirectly (e.g., via eye gaze) tend to dissociate from explicit abilities, which are measured using explicit questions (Onishi & Baillargeon, 2005; Senju, Southgate, White & Frith, 2009). However, these two processes appear closely related in NT adults (Kanske, Böckler, Trautwein & Singer, 2015). Therefore, studying adults facilitates the construction of theoretical models which support an understanding of empathic development (Apperly, Samson & Humphreys, 2009) and identification of diagnostic markers for ASD, clinical and neurodegenerative disorders (e.g., Guastella et al., 2013; Poletti, Enrici & Adenzatio, 2012), and can inform training programmes aimed at improving empathy (for a meta-analysis, see Teding van Berkhout & Malouff, 2016). A challenge for researchers lies in establishing tools sensitive to variation at the upper echelons of the ability, since NT adults tend to perform at ceiling (at or near 100% accuracy) on traditional measures.

2.4 Measuring Mentalising

Mentalising ability can refer to identifying a target's cognitive states (thoughts, beliefs, knowledge and intentions) and affective states (emotions). Both entail setting aside one's own knowledge or feelings in order to attribute states to other agents whose thoughts and feelings

may differ from one's own. Automatic egocentric bias—the tendency for one's own perspective to influence a judgment about a target—has been demonstrated in adults and children in verbal and visual perspective-taking tasks (e.g. Epley, Morowedge & Keysar, 2004), but NT adults show an ability to correct for this (Wang, Miletich, Ramsey & Samson, 2014). Mentalising measures can examine both explicit (deliberate) mentalising and implicit or spontaneous (automatic) mentalising. However, some more elaborative processing, such as belief-attribution, has been shown to be non-automatic in adults (Back & Apperly, 2010).

2.4.1 Implicit (spontaneous) mentalising

The automatic tendency to mentalise about internal states is so ingrained it can apply to non-sentient targets, as exemplified in the following excerpt from Robert Frost's (1969/2013) poem, *A considerable speck (microscopic)*:

A speck that would have been beneath my sight

On any but a paper sheet so white

Set off across what I had written there.

[…]

This was no dust speck by my breathing blown,

But unmistakably a living mite

With inclinations it could call its own.

It paused as with suspicion of my pen,

And then came racing wildly on again

To where my manuscript was not yet dry;

Then paused again and either drank or smelt—

With loathing, for again it turned to fly.

Plainly with an intelligence I dealt.

[…]

It ran with terror and with cunning crept

It faltered: I could see it hesitate;

Then in the middle of the open sheet

Cower down with desperation to accept

Whatever I accorded it of fate. […]

(Frost, 1969/2013, p. 357)

Heider and Simmel (1944) demonstrated this tendency in the laboratory. They showed participants animations of moving shapes and after viewing two triangles moving around, in and out of a rectangular shape (Figure 2), participants tended to talk about the triangles as if they had thoughts and plans. It appeared that participants spontaneously (without explicitly being asked) engaged in mentalising, which led to the attribution of intentional states. This was supported by results from a similar study in which autistic children, who tend to show deficits in mentalising, were found to ascribe fewer mental state terms to the shapes compared to typically developing children (e.g., Abel, Happé & Frith, 2000). Developmental research has adopted similar approaches in order to study mentalising in typically developing children (e.g., Csibra, 2008; Gergely, Nádasdy, Csibra & Bíró, 1995; Hamlin, Wynn & Bloom, 2007); however, investigation of rudiments and shortcomings limits generalisation to the spontaneous attribution of complex, situated mental states with NT adults (although the task has been used in conjunction with brain imaging techniques to examine neural correlates; e.g., Castelli, Happé, Frith & Frith, 2000).
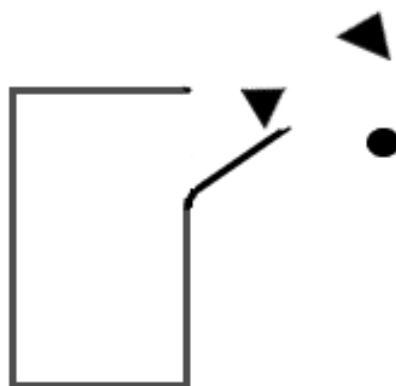
Figure 2. In Heider and Simmel's (1944) classic study, participants viewed a video of animated shapes moving around and within a rectangle. Image: Logeotalexandre (2019), licensed by Creative Commons (CC BY-SA: https://creativecommons.org/licenses/by-sa/4.0/deed.en).

A more recent example of stimuli for testing spontaneous mentalising uses a virtual environment (VE). The Interactive Real World Task (Spiers & Maguire, 2006) requires participants to engage in a driving simulation video game, which involves travelling through London, England, from a ground-level first-person perspective. The creators of the game (Sony Computer Entertainment) produced a bustling cityscape using real photographs of the streets of London and computer-generated passers-by. The driveable streets were generated using Ordinance Survey map data. After taking part, participants are asked to watch a recording of their simulation and to report, verbally, to an interviewer, what they were thinking during the simulation. The simulation video can then be overlaid with this recording and used to code "ToM events", points at which the participant commented on the thoughts and intentions of themselves or others in the simulation. The original study allowed researchers to observe patterns in functional magnetic resonance imaging (*f*MRI) data in relation to participants' mentalising.

This approach is time-consuming as training is required prior to participation in VEs, though it may provide a worthwhile trade-off for measuring online spontaneous empathic inferencing. VEs offer higher ecological validity compared to other measures (Parsons & Mitchell, 2002), such as viewing videos of animated shapes, although the results are limited to the context of the virtual environment used. Preliminary data from ASD groups indicates that VE-based training can improve mentalising abilities (Kandalaft, Didehbani, Krawczyk, Allen & Chapman, 2013), and so they may be beneficial as a training tool. While the task was designed to elicit spontaneous mentalising, the addition of direct questions and a coding system containing accuracy and complexity variables could be incorporated to facilitate a temporal view of explicit decoding and reasoning processes.

2.4.2 Explicit mentalising: Cognitive content

Self-report mentalising measures provide easy-to-administer indexes of people's own views on their ability and tendency to take the perspectives of others. For example, the perspective-taking dimension of the Interpersonal Reactivity Index (IRI; Davis 1980; which consists of four independent scales probing different empathic processes) measures the self-reported tendency to take others' points of view. Participants indicate agreement with seven statements such as, "I try to look at everybody's side of a disagreement before I make a decision", on Likert scales. The approach relies on the participants' abilities to accurately reflect on their mentalising, and risks socially desirable responding (where participants' answers are biased in socially desirable directions, e.g., by reporting a higher tendency towards an agreeable social behaviour than they actually possess).

Response bias is less of an issue when using behavioural approaches. False belief tasks have been widely used in tests of mentalising. They probe the ability to set aside one's own knowledge or "true beliefs" in order to identify a target's mistaken beliefs (Dennett, 1978). The traditional false belief task (Wimmer & Perner, 1983) depicts belief-states through social vignettes. In the conventional object-transfer paradigm, participants must identify a target agent's mistaken belief about the location of an object through understanding that the agent lacks the knowledge that the object has been moved. This can vary in complexity, for example, person A wrongly believes that the sweets are in the opaque jar because they did not witness person B move them to the cupboard (first-order); person B wrongly believes that person A will look for the sweets in the jar, unaware that person A secretly watched them being moved (second-order).

False belief tasks have been used in child development studies (for a meta-analysis, see Wellman, Cross & Watson, 2001), ASD (Baron-Cohen, Leslie & Frith, 1985), psychiatric disorders (Frith & Corcoran, 1996), brain damage (Winner, Brownell, Happé, Blum & Pincus,

1998), stroke (Happé, Brownell & Winner, 1999), and Alzheimer's (Le Bouc, Lenfant, Delbeuck, Ravasi, Lebert, Semah & Pasquier, 2012). As children typically pass first- and second-order tasks aged 4-5 (Astington & Dack, 2008) and 6-7 (Perner & Wimmer, 1985) respectively, these measures tend to show ceiling effects with adults. Adaptations for use with NT adults include a version where participants rate the likelihood that a protagonist will look for her violin in various locations (Birch & Bloom, 2007). In this version of the classic task, participants are privy to the object's location in two conditions, which are varied by plausibility (plausible and implausible). In the plausible condition, participants are told that, following the protagonist's exit, her violin has been moved to a different container. The room has been rearranged, and the violin in the new container has been moved to the same location as the old container (where the protagonist might plausibly look). In the implausible condition, the violin had been moved to a new container and positioned in a new location, and so it would be implausible for the protagonist to find it there. Participants are asked to identify the percentage likelihood that the protagonist would look in each location. This test was shown to be sensitive to the interference of NT adult participants' knowledge ("reality bias"; Mitchell, Robinson, Isaacs & Nye, 1996).

False belief understanding has become synonymous with ToM; however, the construct validity of false belief tasks has been called into question (e.g. Bloom & German, 2000). For example, the False-Belief Localizer tool for isolating the neural basis of false belief representation (Dodell-Feder, Koster-Hale, Bedny & Saxe, 2011; Saxe & Kanwisher, 2003), is often referred to as the "ToM Localizer", yet the neural pattern diverges from meta-analytic accounts of the ToM network (Spunt & Adolphs, 2014). In developmental populations, poor false belief task performance may reflect general task demands such as verbal or memory demands (Siegal & Beattie, 1991; Sullivan, Zaitchik & Tager-Flusberg, 1994). Some individuals with ASD pass second-order tasks while exhibiting real-world social cognitive

difficulties (Happé, 1994), suggesting that they may recruit compensatory verbal strategies (Happé, 1995) such as knowledge of complement syntax (Lind & Bowler, 2009) in order to pass the task. Social animation tasks such as those used in tests of implicit or spontaneous mentalising (e.g., Castelli et al., 2000; Heider & Simmel, 1944) circumvent this issue, though they lack the range of epistemological and emotional information present in real-world scenarios. Instead, naturalistic narrative stimuli enable researchers to embed target states in specific contexts (e.g. Frith & Corcoran, 1996; Saxe & Wexler, 2005), which may inhibit the use of other strategies (Happé, 1995). Happé's (1994) Strange Stories Task assesses comprehension of short, naturalistic narratives, including joke, lie, appearance/reality, and contrary emotions. The range of narratives have proven more sensitive to subtle between-group differences than false belief tasks, paving the way for more complex narrative-based approaches (see Section 2.8).

2.4.3 Explicit mentalising: Affective content

The ability to accurately infer affective states is often tested using facial expressions as stimuli because the ability to recognise emotions precedes mentalising (Mitchell & Phillips, 2015). Emotion recognition tests traditionally require participants to identify basic emotions (happiness, sadness, anger, fear, surprise, contempt, and disgust; Ekman & Friesen, 1971)[3] presented in photographs or brief video-clips of posed facial expressions. The emotion perception literature has primarily focused on macroexpressions: full-face unconcealed expressions lasting more than 0.5 seconds; however, strong agreement of the basic emotions can result in ceiling effects with NT adults (in other words, given enough time, adults tend to find identifying "happy", "angry" and "sad" faces easy). One approach is to speed up presentation so that stimuli represent microexpressions (Ekman & Friesen, 1976), which last

---

[3] See Awasthi and Mandal (2015) for a critical review of existing theoretical models of emotional facial expressions (sociocultural, biological, computational) and their proposal of a comprehensive, multidisciplinary model.

up to 0.25 seconds and are usually fragmentary (appearing on the top or bottom half of the face). A challenge for this method is that even brief presentations can remain on the retina for longer (e.g., Brief Emotion Recognition Test, BART; Ekman & Friesen, 1974), though this is resolvable by incorporating neutral expressions as forward-backward masks (appearing before or after the critical image; Matsumoto et al., 2000). Microexpressions are involuntary, tending to signal concealed or altered emotional expressions, so perceiving them is likely to reflect the advanced capacity to detect deception in real-life interactions (Frank & Svetieva, 2015).

The Reading the Mind in the Eyes Test (RMET; Baron-Cohen, Wheelwright, Hill, Raste & Plumb, 2001) is a widely used test of mentalising. It requires participants to attribute the most appropriate mental state term (e.g., "ashamed", "nervous", "suspicious", "indecisive") to photographs of the eye-regions of faces. The task probes non-automatic processes (Bull, Phillips & Conway, 2008) and was designed to detect subtle deficits (Baron-Cohen, Jolliffe, Mortimore & Robertson, 1997). It has been applied to a range of domains including brain studies (Adolphs, Baron-Cohen & Tranel, 2002), dementia (Gregory, Lough, Stone, Erzinclioglu, Martin, Baron-Cohen & Hodges, 2002) and clinical disorders (e.g. Fett, Viechtbauer, Dominguez, Penn, van Os & Krabbendam, 2011). The RMET demonstrates particularly strong predictive power with ASD groups, supporting its validity as a measure of the mentalising deficit associated with ASD. In Baron-Cohen, Wheelwright, Hill et al.'s (2001) original study, performance negatively correlated with the Autism Spectrum Quotient (50 forced-choice items about social skills, communication skills, imagination, attention to detail, and tolerance of change; Baron-Cohen, Wheelwright, Skinner, Martin & Clubley, 2001). This finding may reflect the purity of the stimuli, which minimises the opportunity to depend on alternative cues (such as verbal cues; see Happé, 1995).

Although the RMET has been used in several studies with NT adult participants (e.g., Kidd & Castano, 2013; Mar et al., 2006), it has been found not to effectively discriminate

between individuals exhibiting high levels of mentalising ability (Black, 2018). Furthermore, recent research has indicated that it is subject to bias because it contains only Caucasian actors and is heavily reliant on participants' vocabularies, requiring participants to select between linguistic mental state classifiers. Dodell-Feder, Ressler and Germine (2020) found more pronounced effects of education, race and ethnicity on RMET scores in comparison to other tasks (emotion identification, discrimination and non-verbal processing speed tasks). Oakley, Brewer, Bird and Catmur (2016) cautioned that the RMET provides an index of emotion recognition ability rather than the more complex mentalising processes. This is an important distinction as emotion recognition and other processes (e.g., attribution of thoughts and beliefs via explicit reasoning) can dissociate (Oakley et al., 2016). Emotion recognition stimuli processed via a single modality (including facial/body images or auditory voice recordings, e.g., Rutherford, Baron-Cohen & Wheelwright, 2002) present a general issue of ecological validity, and a specific problem for studies with NT adults. Older adults tend to perform poorly in comparison with young adults on static emotion recognition tests but outperform them at recognising continuous emotions in dyadic interactions (Sze, Goodkind, Gyurak & Levenson, 2012). Dynamic stimuli can be used to avoid problems associated with using simple static images (Biele & Grabowska, 2006; Halberstadt, Dennis & Hess, 2011), although both static and dynamic stimuli lack information about contexts in which real-world scenarios are embedded (Achim, Guitton, Jackson, Boutin & Monetta, 2013).

2.4.4 Explicit mentalising: Cognitive and affective content

Developments in approaches to measurement have led to more complex mentalising tasks that yield data pertaining to both cognitive and affective mental content, at different levels of complexity. This facilitates the indexing of ability between processes and within participants. The computerised Yoni test (Shamay-Tsoory & Aharon-Peretz, 2007) is one such example. It requires integration of visual and verbal cues and has the capacity to

generate both behavioural and neuroimaging data. A series of vignettes feature a central character, "Yoni", depicted by a schematic face—a simple cartoon "smiley"—and four images of a single category (e.g., faces, animals, fruit or transport) alongside sentences containing blanks (Figure 3). Participants indicate by selecting (clicking with the cursor) the appropriate image based on what Yoni is close to, thinks about, loves, does not love, or identifies with (first-order); and whose misfortune Yoni gloats over, whose success Yoni envies, and items Yoni thinks about, has or loves, that another character thinks about, has or loves (second-order). The task entails interpretation of proximity, eye-gaze and facial expressions, and provides measures of accuracy and response time across cognitive, affective and physical (control) trials.
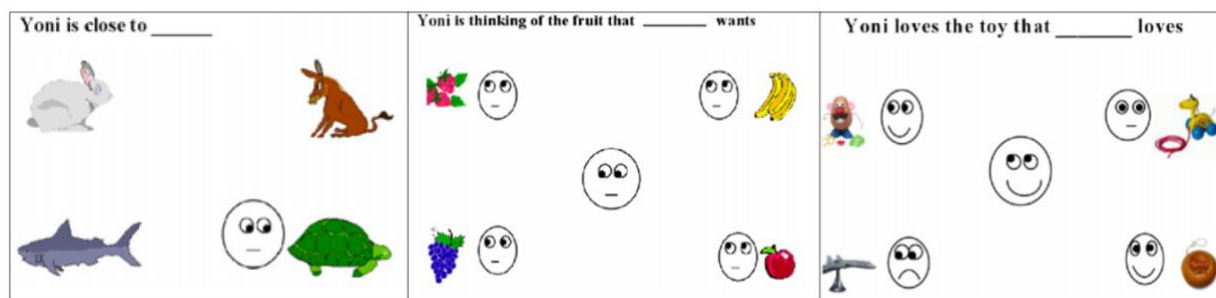
Figure 3. Yoni test stimuli. Example first-order physical (control) trial (left), second-order cognitive trial (centre) and second-order affective trial (right). Adapted from *Neuropsychologia, 45,* Shamay-Tsoory, S. G., & Aharon-Perez, J., "Dissociable prefrontal networks for cognitive and affective theory of mind: A lesion study", 3054-3067. Copyright Elsevier (2007). Reprinted with permission from Elsevier.

In the original study, success was higher for affective content compared to cognitive content trials, a finding replicated by Kalbe et al. (2010), who suggested that additional facial expression cues in the affective condition facilitated decision-making (however, the scoring system does not separate out the emotion recognition dimension from the other affective content). Nonetheless, second-order differences between controls and patients with ventromedial frontal lobe damage were observed only in the affective condition, indicating that the task was sensitive to dissociation between cognitive and affective neural systems (Shamay-Tsoory & Aharon-Peretz, 2007). The Yoni test has shown sensitivity to variation in

NT adults where false belief tasks have proven insufficient (e.g., Kidd & Castano's, 2013, study of fiction effects on empathic accuracy); however, it is possible that the simplistic stimuli enable participants to form basic object-agent associations rather than engage in mentalising *per se* (this is also a criticism of false belief tasks; Perner & Ruffman, 2005). In contrast, the Why/How Task (Spunt & Adolphs, 2014) prevents the formation of basic associations by asking participants "how" (physical) and "why" (mentalising) questions about human behaviours depicted in photographs. It was designed for *f*MRI studies but is also capable of providing reliable behavioural (accuracy and response time) data.

Stimuli comprised of simple social images have presented opportunities to examine the brain basis for behavioural differences between participants. Although their central limitation is a lack of ecological validity—they do not reflect the complexity of real-world social situations—they do enable specific aspects of mentalising to be isolated, though this raises the issue that some participants may utilise alternative, unmeasured strategies.

2.5 Measuring Experience-sharing

Self-report measures can probe trait experience-sharing tendencies with real or fictional others. The Emotional Contagion Scale (Doherty, 1997) is a 15-item measure of the tendency to "catch" five basic emotions: love, happiness, fear, anger, and sadness. Participants indicate agreement with statements like, "I get filled with sorrow when people talk about the death of their loved ones," via a Likert scale anchored with "never" and "always". The questions pertain to real-world others, except for one item about feelings towards fictional scenarios ("I cry at sad movies").

Emotion contagion can emerge from simple exposure to a target's emotions or from interaction with a target. Guillory, Spiegel, Drislane, Weiss, Donner and Hancock (2011) took an experimental approach to manipulating affect and examined emotion contagion in interactive groups. Participants were assigned to groups of three and asked to watch film

clips. In the negative condition, one participant (the "experiencer") watched a clip that reliably induces negative emotions while the other two participants watched a neutral clip. In the neutral condition all three participants watched the neutral film clip. After viewing the film clips all participants filled in the Positive and Negative Affect Schedule (PANAS; Watson, Clark & Tellegen, 1988), a widely used measure of state emotion. Next, in their groups, participants were asked to chat, to solve word scrambles, and to listen to music, ostensibly for a test of multi-tasking ability. The "experiencer" in the negative group was assigned a word scramble task which used negative primes (the solutions were negative words such as "malice") while listening to abrasive heavy metal music. This procedure was designed to maintain the negative emotion induced by the film. All other participants (neutral group participants and the partner participants in the negative group) listened to gentle jazz music and solved word scrambles that yielded neutral words. Finally, participants completed another questionnaire measure of state emotion (the affect items from the Circumplex Affect Scale; Russell, 1980) and were asked questions about group members' emotions.

The results of Guillory et al. (2011) revealed that tension was contagious: partners of the experiencer (whose negative affect was induced via the tasks) reported feeling significantly more tense compared to the neutral group. However, other emotions were not found to be contagious. This finding may have been due to a central limitation of the design: the lack of control over the "experiencer's" expressiveness (e.g., they may have attempted to conceal their negative emotions from their groups), which would need to be addressed before drawing conclusions about the transferability of specific emotions in this context. To overcome this issue, a confederate actor trained to display a particular mood state can be planted as a "participant". An example of this comes from Barsade (2002) who measured participants' affect before and after a group interaction (his was a 2×2 design varying the confederate's pleasantness and level of energy). He used external raters as well as

questionnaires yielding both self-report and coded measures of emotion contagion (the findings from this study were that participants who were with the pleasant confederate showed an increase in pleasant mood over time and, interestingly, the low energy groups also became more positive over time). Group tasks like these benefit from the ecological validity of the setting but are resource heavy, and their interactive nature makes it difficult to isolate the specific cues (e.g., facial or verbal) responsible for any effects.

Emotion contagion does not necessarily take place face-to-face between people. Emotions can be transmitted through cultural artefacts such as music, film and images (Giuliana & Carvalho, 2016). Therefore, the vicarious experiencing of emotions can be assessed using music, pictorial or film stimuli. Music appears to elicit the physiological and experiential aspects of the emotional response system (Lundqvist, Carlsson, Hilmersson & Juslin, 2009), and participants who watched films of a target describing a significant happy or sad life event tended to experience the same emotions as the target (Hsee, Hatfield, Carlson & Chemtob, 1990). Such approaches tend to trade ecological validity for higher levels of control; however, dynamic stimuli (e.g., video clips) may offer a reasonable compromise as they are more naturalistic than static pictorial stimuli but offer greater control than interactive group situations.

Self-report measures can assess the tendency to share in the emotions of fictional characters. For example, the 7-item fantasy subscale of the IRI (Davis, 1980) assesses the tendency to identify with fictional people and situations, with all but one item ("I daydream and fantasise, with some regularity, about things that might happen to me") pertaining to absorption in narrative worlds and empathy for fictional characters (e.g., "after seeing a play or movie, I have felt as though I were one of the characters"). It has been suggested, therefore, that fantasy scores provide an index of imaginative abilities (Baron-Cohen & Wheelwright, 2004) and that the scale can be treated as the trait tendency to become immersed in stories

(e.g., Mar et al., 2006; see also Hall & Bracken, 2011). The ability to immerse oneself in a storyworld is akin to the Einfühlung concept (projecting oneself into another object or person and experiencing it from within); however, how far this particular scale represents cognitive imaginative versus affective empathic processes, has been debated (Jordan, Amir & Bloom, 2016).

Because empathy for fictional characters appears to relate to empathy for real-world others (Nomura & Akai, 2012), empathy can be measured via the tendency to share in the emotions of fictional characters using questionnaires following a reading task. For example, The Identification with Character scale (Iguarta, 2010) is a 14-item scale comprising two dimensions: "empathic", cognitive and affective reactions to characters (e.g., "I understood the characters' way of acting, thinking or feeling"), and "merging", the experience of becoming a character and losing of self-awareness (e.g., "I had the impression I was really experiencing the story of the characters"). Agreement with statements is indicated on a Likert scale anchored with "not at all" and "very much". As there remains some debate about whether empathy requires the self and other to remain distinct (i.e., a lack of merging; Cuff et al., 2016), the dimensions of this scale could be applied depending on the theoretical standpoint.

2.6 Measuring Prosocial Concern

Prosocial concern, sympathy or "empathic concern" (Batson, Fultz & Schoenrade, 1987; Davis, 1980) is the desire to help alleviate another's suffering. The emotion experienced should be incongruent with that of the target (the observer does not share in the distress), but appropriate to the situation (e.g., Baron-Cohen & Wheelwright, 2004). In this context, "appropriate" refers to feelings of sympathy and compassion rather than, for example, Schadenfreude (pleasure derived from another's misfortune; Zickfeld et al., 2017). Unlike mentalising and experience-sharing constructs, with the empathic concern construct

there is a prerequisite for an "unfortunate" target. The target's experience must be negatively valenced (e.g., distress) in order to activate the desire to alleviate their misfortune and improve their welfare. Thus, the concern component is invoked by an awareness of need (recall that the empathy concept was developed alongside Hume's writings on sympathy, and that "pathos" constitutes part of the root of the term).

The empathic concern subscale of the IRI is the most widely used measure of concern. It indexes the self-reported tendency to experience sympathy or concern for others (e.g., "I often have tender, concerned feelings for people less fortunate than me"). Participants indicate agreement with seven statements on Likert scales. As with other self-report approaches, the measure may be susceptible to upward bias caused by socially desirable responding. A different self-report approach, the Empathy Index (Batson et al., 1987), asks participants to indicate the extent to which they have felt six emotions ("soft-hearted", "compassionate", "moved", "tender", "warm" and "sympathetic") towards the stories of others (targets can be real or fictional), and Likert scales are anchored with "not at all" and "extremely". While these two approaches do not necessarily map onto real-world empathising, they do provide an index of the extent that a person tends to become emotionally affected, and they have been shown to correlate with other (mentalising and experience sharing) dimensions of the IRI.

2.7 Measuring Prosocial Behaviour

Prosocial, helpful or altruistic behaviours can also be measured using self-report. For example, the Self-report Altruism Scale (Rushton, Chrisjohn & Fekken, 1981) requires participants to indicate the frequency with which they have carried out 20 acts that refer to helping acquaintances, strangers, and charities (e.g., "I have given money to a stranger who needed it or asked me for it") on a Likert scale from "never" to "very often". Penner et al. (1995) produced a prosocial personality battery comprising 56 items across two factors: other-oriented empathy (which includes 19 items from the IRI), and helpfulness. However, socially

desirable responding represents a potential limitation, and some items concerning specific behaviours (e.g., donating blood or giving money to charity) may be impacted by factors beyond prosociality (such as eligibility or financial status).

Task-based prosocial behaviour measures include dictator games in which participants decide how to distribute cash sums between themselves and other participants or, similarly, prosocial orientation tasks where they distribute points between themselves and other players (for an overview see Camerer, 2003). In these tasks, participant behaviour is measured in terms of altruistic and egoistic choices. For example, Koopman (2015) measured altruism dichotomously: "yes" if participants gave their participation fee to charity and "no" if they declined to do so. The Pen Drop task (van Baaren, Holland, Kawakami & van Knippenberg, 2004) also yields a dichotomous outcome measure. While collecting in the experimental materials, the experimenter drops a set of pens, ostensibly by accident: what is measured here, simply, is whether or not the participant shows prosocial behaviour by lending a hand in picking them up. The approach can also incorporate different objects (e.g., the experimenter drops a pile of books; Macrae & Johnston, 1998). A drawback of this latter task is that as it requires live participation in the lab and consistent behaviour from the experimenter, there may be a substantial experimenter effect (e.g., Does, Ellemers, Dovido, Norman, Mentovich, van der Lee & Goff, 2018; Rosenthal, 1976). In contrast, prosocial orientation and dictator games, or charity giving measures, can be adapted for presentation via a computer (online or in the lab) which can support a more standardised approach.

2.8 Multidimensional Approaches

2.8.1 Fiction, film and interactive stimuli

Using complex fictional narratives allows for the inclusion of a range of contextual variables. They provide opportunities for measuring both mentalising and experience-sharing strategies using explicit and implicit mental state questions. For example, participants in the

Short Story Task (SST; Dodell-Feder et al., 2013) read a fictional story about two characters whose romantic relationship breaks down (Hemingway, 1988). The text contains first- and second-order mental states and interpretation of them requires the synthesis of contextual, verbal and physical information presented in the text. Semi-structured questions probe explicit and spontaneous mentalising with explicit items scored from 0-2, and a single spontaneous question as a dichotomous yes/no variable. However, as the spontaneous question prompts participants to provide "the character's thoughts, feelings and intentions when it applies to the question" (Dodell-Feder et al., 2013, p. 4), the implicit versus explicit distinction is not clear-cut. The coding scheme does not distinguish cognitive and affective, or first- and second-order attributions (indicated through low internal consistency; $\alpha = .54$), and so the scoring system could be developed in the future to support a more nuanced picture of empathic accuracy (see Dodell-Feder et al., 2013, for some recommendations). In Dodell-Feder et al.'s original study, the SST demonstrated sensitivity to variation among NT adults (with scores ranging from 2-14 out of 16 available points), and concurrent validity with mentalising and experience-sharing measures (the RMET and the IRI fantasy subscale).

Using film stimuli enables researchers to present verbal, facial and bodily cues within dynamic dyadic and group interactions (e.g. Barnes, Lombardo, Wheelwright & Baron-Cohen, 2009; Bazin, Brunet-Gouet, Bourdet, Kayser, Falissard, Hardy-Baylé & Passerieux, 2009; Golan, Baron-Cohen & Golan, 2008). They also provide the opportunity to use actors, which offers increased control over the content of the interactions presented. For example, the Interpersonal Perception Task (IPT; Constanzo & Archer, 1993) consists of thirty short, videotaped scenes, though a shortened version containing fifteen scenes is available (IPT-15). Each scene presents one of five types of social interaction: status, intimacy, kinship, cooperation and deception. Each scene is accompanied by a single multiple-choice question that requires participants to decode an aspect of the interaction, such as, "what is the

relationship between the man and the woman?" Although the IPT-15 has been shown not to correlate with some social ability tasks (Lindvall, 2008), it has been found to correlate with peer ratings of people's social skills (Constanzo & Archer, 1993). This indicates that task scores converge on some real-world outcomes, but highlights heterogeneity among social ability tasks.

The Movie Assessment of Social Cognition (MASC; Dziobek et al., 2006) presents more complex social scenes, which could enable a range of empathic dimensions to be examined. The film features four characters at a dinner party, where prominent themes are romance and friendship. A script development process (Field, Meyer & Witte, 2001) generated realistic characters that display both stable traits and transient states. Participants are asked direct questions about the characters' cognitive and affective mental states, which requires the interpretation of vocal, physical and contextual information, alongside classic concepts such as false beliefs, metaphor and *faux pas*. Versions of the MASC include the original German and dubbed English editions (dubbing did not interfere with participants' task focus, Dziobek et al., 2006, and generally does not impact information processing, Koolstra, Peeters & Spinhof, 2002).

Dziobek et al. (2006) found that the MASC converged with three extant mentalising measures: a basic emotion recognition task, the RMET, and the Strange Stories Task (shortened). Specifically, MASC scores predicted Strange Stories Task performance in participants with Asperger Syndrome, and emotion recognition in controls, which indicated that verbal strategies may have compensated for facial processing difficulties. This highlights the trade-off between greater ecological validity and the problem of identifying specific processes used in complex tasks. Control questions are useful for partialling out variance caused by more general abilities; however, both groups (participants with Asperger Syndrome and NT participants) performed at ceiling on the MASC control questions, and so more

challenging questions would be required to account for other cognitive abilities (Dziobek et al., 2006; Heavey, Phillips, Baron-Cohen & Rutter, 2000). Notably, the MASC was more sensitive to group differences than the established measures it was compared against, and this was supported by a more recent finding that participants with ASD showed impaired MASC, but not RMET performance, when compared to participants with alexithymia (a condition characterised by impaired emotion recognition that often co-occurs with ASD; Oakley et al., 2016).

Generalisability and longevity of film tasks may be limited due to the contextually specific nature of mental state inferences (for example, interactions may be better understood by similar age-groups to the characters; Griffiths, 1997). This is also true, to some extent, for fictional prose, where character names and the language used in dialogue may signal the ages of characters or the social era in which the story is situated. However, this information is more easily updated in prose stimuli, whereas film's additional visual cues such as clothing and objects are more difficult to alter (consider how easily a film can be dated by hairstyles, fashion, or mobile phones used by its characters). The impact of identification with a particular social context is especially problematic for the study of empathic processes for which there may be an ingroup advantage (e.g., Matsumoto, Olide & Willingham, 2009).

An approach that overcomes this problem is the use of social scenarios that participants are actively involved in. After all, individuals do not only observe the social world, they interact with it too. Interactive approaches to measuring empathic accuracy include a participative version of the Empathic Accuracy Paradigm (Ickes, Stinson, Bissonnette & Garcia, 1990) where pairs of participants are covertly filmed waiting to participate in an experiment. After being debriefed, participants watch the footage back individually, and they are asked to describe their own thoughts and feelings, as well as to infer those of their partner. Partner inferences are scored for accuracy, which constitutes the dependent variable. The

procedure is socially valid—the interaction is real—but contingent on individuals accurately articulating their own mental states (Cuff et al., 2016). It is also limited to the range of states naturally occurring in the context, and participant motivations may influence accuracy: for example, interest in and perceived attractiveness of the partner factored in the original study. The original study did not account for the extent that accuracy may evolve or fluctuate over the course of the interactions, but because target moments are time-logged, this temporal aspect could be examined in future designs.

2.8.2 Mentalising and experience-sharing stimuli

The measures described in this chapter tap into different components of empathy. Smith (2017) invoked Goldman's (2011) assertion that there are different "routes" to empathy. He pointed out that these different capacities do not themselves represent empathy, rather they reflect processes that can lead to empathic accuracy, a state of understanding of a target:

> None of the psychological phenomena such as imitation, emotional contagion and perspective-taking are to be identified with empathy. Nor are any of them strictly necessary for empathy […] Nevertheless, as I have indicated, these various psychological phenomena can feed into empathy […] On the account that I have sketched, empathy is not a process of any sort, rather it is a state in which one arrives having undergone those grounding processes, whatever they may have been. (Smith, 2017, p. 718)

As discussed, mentalising and experience-sharing processes represent routes to the state of empathy, and their efficacy can be measured using empathic accuracy toolkits such as those described in this chapter. Due to their dissociable neural correlates, mentalising and experience-sharing have generally been studied separately (although recent research has indicated some overlap between these processes; Zaki & Ochsner, 2012). While multidimensional approaches (e.g., Davis, 1980) have the capacity to measure and contrast both components, they have yet to be operationalised in this way.

Zhou et al. (2017) produced a behavioural paradigm with which to index and compare two pathways to empathic accuracy. They asked participants to estimate the emotional ratings that target individuals ("experiencers") had given in response to positive, negative and neutral photographs. Participants were assigned to either the "theorisation" strategy, where they watched short videos of the experiencers' dynamic facial responses to the photographs, a "simulation" strategy, where they viewed the same photographs as the "experiencers" and were able to use their own reactions as proxies, or a "simultaneous" condition where the photographs and the videos were presented side-by-side (Experiments 1-4; Figure 4). Consequently, the theorisation and simulation stimuli reflect mentalising (emotion recognition) and experience-sharing (using one's own emotion as proxy) exercises, respectively. A novel mapping of these conditions onto the mentalising and experience-sharing components identified in the model outlined in section 2.2, provides a framework through which to examine mentalising and experience-sharing sub-processes, and an opportunity to compare their respective value for empathic accuracy.
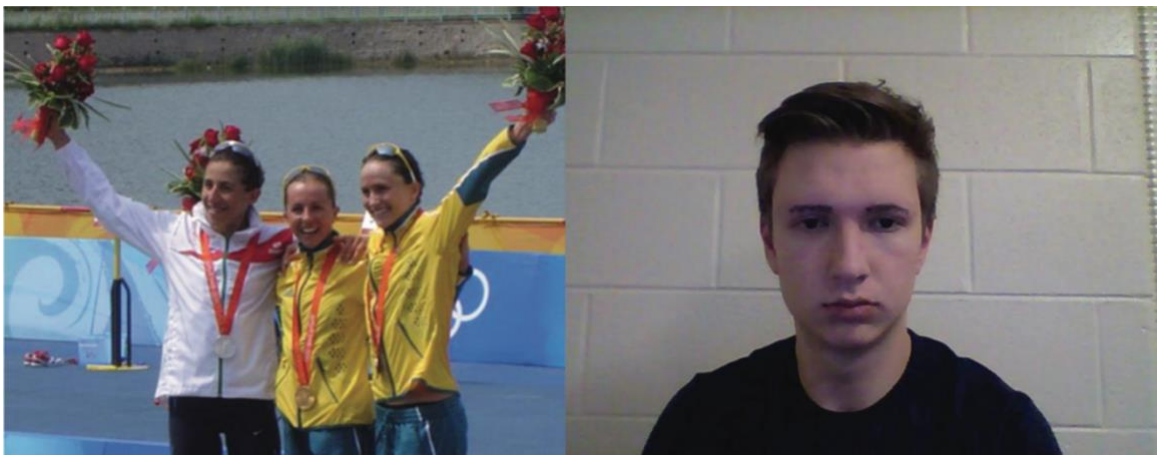
Figure 4. Illustrative example of Zhou et al.'s (2017) simultaneous condition showing a photograph viewed by the experiencer (left) and video of the experiencer viewing the photograph (right). In the theorisation condition the photograph would be greyed out, and in the simulation condition the video would be greyed out. Reprinted from Zhou et al. (2017) and licensed under the GNU Free Documentation Licence (https://commons.wikimedia.org/wiki/Commons:GNU_Free_Documentation_License,_version_1.2).

In Zhou et al.'s study, the two strategies for interpreting mental states were shown to be unequal, with the simulation strategy leading to higher empathic accuracy compared to the theorisation strategy. In Experiment 1, for example, participants in the simulation condition were more accurate than participants in the theorisation condition 92% of the time and participating in the simultaneous condition did not show an increase in accuracy beyond simulation. This paradigm enables testing of either mentalising, experience-sharing or both routes to empathic accuracy, via one experiment. Future research could use it to establish how far mentalising and experience-sharing impact concern and prosocial behaviour. While the task itself provides a behavioural measure, it is reliant on the accuracy of the targets' self-reports. Furthermore, the authors acknowledged that "simulation" may not represent the most effective strategy across all contexts. Situational factors, such as the valence of the target emotion (Kauschke, Bahn, Vesker & Schwarzer, 2019), similarity to the target and variation across people's experiences (e.g., Barrett, Mesquita & Gendron, 2011) would likely modulate the effect of strategy.

2.9 Selecting and Combining Measures

Extant research on empathic accuracy has primarily focused on children, ASD, and clinical populations, and standard measures can suffer ceiling effects with NT adults. However, several classic tests have been adapted for research with adult cohorts. This chapter has described a range of paradigms used to investigate aspects of empathy in NT adults, some of which have been employed in the present research.

Tasks measuring specific dimensions such as facial emotion recognition and cognitive mentalising provide "pure" measures that can distinguish between first-order and higher level processes, and can incorporate both cognitive and affective dimensions (e.g., Yoni test, used in the research presented in Chapter 7). However, these approaches can risk subtracting out key processes, particularly as real-world social displays can be suppressed, and so interpreting

mental states may require integration of verbal, physical and contextual information (McDonald, Flanagan, Rollins & Kinch, 2003). In contrast, interactive tasks, and those using VEs, have proven effective both in social skills development and in studying the neural bases of empathic inferencing. These, however, can be resource-heavy, reliant on participants' introspective abilities (such as when participating as both empathiser and target in dyadic set-ups), and require substantial training.

Less resource-heavy approaches include multidimensional self-report measures, which provide insight into a range of processes and tendencies (e.g., the IRI, employed in the study presented in Chapter 5). Measures such as the IRI have been widely used and are well-validated, though they rely on the ability of participants to accurately reflect on their own thoughts and behaviours and report them consistently. Using fiction as stimuli (e.g., the SST, used in the research presented in Chapter 7), provides realistic, complex scenarios through which to assess different dimensions of empathy. Fictional film enables emotion recognition to be tested alongside more complex processes, but the visual cues available in film may impact engagement with characters (and updated remakes are resource-heavy due to the production of script, employment of actors, crew and locations). Fictional prose stimuli are less resource-heavy and relatively straightforward to translate and reproduce compared to film. Questions can be designed to probe a range of empathic processes (thus reflecting the multiplicity of the empathy construct and its multifaceted cognitive architecture; Christensen & Michael, 2016). However, they do not allow the ability to use facial or bodily cues in empathic inferencing (apart for those that may be described in the text) to be tested, and so an additional task would be required to measure this ability. Zhou et al.'s (2017) paradigm provides one such opportunity, and it enables measurement of both mentalising and experience-sharing strategies. This paradigm is used in the research presented in Chapter 6.

Multidimensional measures can incorporate complex scoring systems in order to avoid compensatory strategies masking selective deficits when being used as diagnostic tools, and they provide opportunities to probe variation both within and between NT adult participants. In this way, multidimensional stimuli need not problematise construct validity, but they could prove fruitful to a nuanced assessment of the range of processes involved in empathy. The concomitant use of multidimensional and simpler tasks has enabled researchers to test the validity of multidimensional measures of empathy. It also enables the dimensions of empathy that may be enhanced through engagement with particular stimuli, to be clarified. This approach was taken in the research presented in Chapter 7, which measured empathic accuracy using both the Yoni test and the SST.

2.10 Summary and Moving Forward

This chapter defined the empathy construct and provided a rationale for studying empathic abilities in NT adults. A range of measurement paradigms were described, and it was proposed that a multidimensional approach would facilitate a nuanced understanding of empathic abilities in NT adults. Mentalising and experience-sharing represent routes to empathic accuracy and, consequently, to prosocial concern and prosocial behaviour. Because these skills develop through the lifespan and can be improved through training, a growing field of research has started to examine the methods through which they may be enhanced. The next chapter will turn to one activity that has been theorised to support the development of empathic abilities: engagement with fiction. The field of research examining fiction effects on empathy has incorporated a range of unidimensional and multidimensional empathy measures. The review of empathy measurement paradigms presented in the current chapter was designed contribute to a more nuanced understanding of their results.

Chapter 3: Fictional Worlds, Real Minds

The previous chapter provided examples of paradigms used to test a range of empathic abilities, highlighting the multidimensionality of the empathy construct. A model of empathy was developed from Zaki and Ochsner's (2012) tricomponent approach. This distinguished between mentalising and experience-sharing strategies for empathic accuracy, which can, in turn, initiate the prosocial concern and prosocial behaviour components. The chapter established the importance of studying empathic abilities in NT adults, in particular because empathic abilities vary throughout the lifespan, are associated with positive social outcomes and can be augmented through training.

This chapter examines one activity that may enhance adults' empathic abilities: engaging with fictional stories. First, it will define fiction for the purpose of the research and briefly outline the cultural development of fiction-reading for leisure. It will introduce Oatley's (1999) fiction-as-simulation theory of the relationship between fiction and empathic abilities, which constitutes the theoretical framework for the empirical research conducted. It will provide a review of the literature on relationships between fiction and empathic abilities including key studies and recent meta-analyses of correlational and experimental findings. Indices of the strength of evidence will be provided. Correlation coefficients will be reported for correlational effects (with associated significance values and confidence intervals where reported by the study authors). Effect sizes for differences between groups will be converted to Cohen's d where means, standard deviations and group sizes are available, and standardised beta values (β) will be used to report regression coefficients. Otherwise, effect sizes will be provided in their original format (e.g., Hedge's g, z, omega squared [ω^2]; for information on these effect sizes and their comparison to others, see Olejnik & Algina, 2003; Rosenthal, 2019).

With reference to the field of narrative persuasion, the process of becoming immersed in narratives will be identified as a potential factor in fiction effects on empathic abilities. It will be argued that fiction and empathic abilities are both multidimensional and that this complexity may explain inconsistent findings across the literature. The chapter will conclude with a proposal that the multidimensionality of both fiction and empathic abilities should be foregrounded in research aiming to establish the nature of the relationship between them.

3.1 Definition of Fiction

People encounter narratives on a regular basis; they are read, watched, listened to, or spoken in conversation. *Fiction* represents a subcategory of narrative, that which is carefully and purposefully constructed (Oatley, 1999). Fictional narratives are not straightforward to define, as narratives can take many forms. Bruner's (1986) definition of the narrative mode, which "deals in human or human-like intention and action and the vicissitudes and consequences that mark their course" (p. 13), implies a coherent sequence of events surrounding humanlike agents, that could refer to a hugely diverse range of existing compositions (Nünning, 2014). In these terms, "fiction" can include poetry, prose, playscript and some history and biography works (e.g., Djikic & Oatley, 2014; Oatley, 1999).

Drawing on Aristotle's distinction, Oatley (1999) pointed out that history is about what *has* happened, and fiction is about what *could* happen. While nonfiction may contain many of the same features, and elicit similar responses to fiction narratives, it is concerned with information, whereas fiction is concerned with emotion (Oatley, 1999); in other words, fiction's focus is a social one. This distinction appears to matter. Reader responses can differ depending on whether a text is framed as fiction or nonfiction; for example, fiction texts may be read more slowly allowing stylistic features to receive more attention than when they appear in nonfiction texts (Nünning, 2014). A further distinction concerns the construction of character. Fiction authors can provide insight into the private thoughts of unusual or

extraordinary characters, a privilege that is "out of the reach" of journalists (Nünning, 2014, p. 81), historians and biographers.4

The central tenet of the social definition of fiction is that it must feature humans, or actors that bear psychological resemblance to humans (Mar & Oatley, 2008; Oatley, 1999). Isserley, the protagonist in Michel Faber's (2000) *Under the Skin,* is non-human (the reader learns that she is an extra-terrestrial alien whose body has been surgically altered in order that she appear "human"). Yet, the psychological and social themes that she encounters are recognisably "human": gender and power dynamics, elitism, love and compassion. Moreover, stories about non-human objects tend to exhibit recognisably human traits and mental states. In stories with a single human character, such as an explorer isolated on a mountain, the environment often takes the role of another character, imbued with humanlike intentions. The explorer might think, for example, "this mountain is trying to kill me" (Mar, 2009, p. 54).

An attraction to human characters and themes in stories has been demonstrated in research with children. Children have been shown to prefer stories that contain people rather than objects—and the more the better—and that include descriptions of their mental states rather than only their actions (Barnes & Bloom, 2014). Indeed, enduring works of fiction, such as Shakespeare's plays, tend to feature close, interrelated sets of characters that echo the structure and size of the social networks of our species. Not only that, but they tend to exploit people's predisposed biases for information about relationships and status: in dramatic fiction the overarching comedy and tragedy genres culminate in marriage or death respectively, reflecting these themes at their highest stakes (Nettle, 2005). Although many works of fiction contain more than status battles and mating conflicts, "where scientific, political, aesthetic, or technical ideas impinge, they only become truly dramatic to the extent that they are made live

---

4 Simply framing a text as "fiction" or "nonfiction" does not necessarily impact reader responses, rather the quality of the narrative appears crucial (e.g., Green & Brock, 2000; Nünning, 2014). Thus, the connection between the narrative's events and actual events is not important, but its stylistic features and content are.

through the wants and relationships of the characters" (Nettle, 2005, p. 66). In other words, human relationships are central to fiction, and social themes are particularly visible in works that have transcended generations.

Fictional narratives are not only read, they are listened to as radio dramas or audiobooks, viewed onscreen or at the theatre and interacted with via video games, drama and roleplay. The technologies for fiction production continue to evolve. In December 2018, streaming service Netflix released its first interactive film, *Black Mirror: Bandersnatch* (McLean & Slade, 2018): part video game, part choose-your-own adventure book, viewers were invited to make dichotomous yes/no decisions at key plot points, ultimately leading to five possible concrete endings, and multiple ways to get there. Due to the increasing range of technologies that mediate engagement with particular stories, people can be avid consumers of fiction without necessarily identifying as "readers". At the extreme, a person need not have read a single book to be familiar with multiple works of contemporary and classical literature. Instead, they may have watched TV or film adaptations, listened to radio plays or seen versions performed onstage. The definition of fiction cannot be constrained to a particular medium. Therefore, fiction is defined, in line with Oatley (1999), as *purposefully constructed text-based narrative that features agents and their interactions, presented via any channel that mediates comprehension of that narrative.*

This definition does not limit the research to full story arcs, which is in line with extant studies that have employed as stimuli both complete stories (e.g., Kidd & Castano, 2013) and extracts from longer stories (e.g., Bal & Veltkamp, 2013). The specification "text-based" denotes the exclusion of other forms of art such as painting, sculpture, and live art, although these forms may contain storytelling elements. There are several researchers already conducting fascinating studies in visual aesthetics, emotion and empathy (e.g., Freedberg & Gallese, 2007; see also Currie, 2011), and future dialogue between these areas of study may

converge to form a broader account of narrative impact across art forms. Text-based fictional narratives uniquely provide opportunities to engage with elaborate descriptions or re-enactments of complex social situations that closely imitate their real-world counterparts.

3.2 Fiction and Social Life

> Lord! When you sell a man a book you don't sell him just twelve ounces of paper and ink and glue—you sell him a whole new life. Love and friendship and humour and ships at sea by night—there's all heaven and earth in a book, in a real book I mean. (Morley, 1917/2010, p. 43)

Telling stories is culturally universal (Brown, 1991) and has existed since humans developed the capacity for speech (e.g., Zipes, 2012). Contemporary hunter-gatherer societies have strong oral storytelling traditions which communicate warnings and social norms and enable cooperation and coordination (Smith et al., 2017). The *Cosmic Hunt* story tells of a bear being chased by three hunters. On being hit, the bear bleeds over the leaves of the forest floor, creating the colours of autumn. It finally escapes its pursuers by climbing up a mountain and into the sky, where it emerges as the constellation Ursa Major. Versions of this myth have been found in Ancient Greece, Siberia, northern Europe, and the Americas, which indicates that it was being spoken when there was a land bridge connecting the land areas that are now Alaska and Russia. This dates the first oral tellings of the story to between 28,000-13,000 BC (Storr, 2019).

The practice of recording narratives, from cave paintings (such as a volcanic eruption depicted in the Chauvet cave in France dating back 36000 years; Nomade et al., 2016) to contemporary books, film and interactive media, allowed original versions of stories to be communicated to multiple people over time. The first printed story was the *Epic of Gilgamesh*, for which constituent sections were carved onto stone tablets as early as 2100 BC (the first surviving written version is from around 1800 BC; Dalley, 2000/1989). Since, it has inspired many contemporary works of literature, music, theatre, film, TV and video games

(Ziolkowski, 2011). Similarly, Aesop's Fables were originally spoken and only written down in the 200s BC—three centuries after his lifetime—and several major texts including the Bible, the Sanskrit poems and the Mayan Codices were probably originally communicated orally. Like Aesop's Fables, they remain in print millennia later.

Before the invention of the printing press in the mid 1400s, all important texts had to be handwritten, which was a costly exercise. With the development of mechanical printing came a sudden increase in the publication and distribution of books, and by the late 18th century, reading was widespread.5 This led to a proliferation of libraries in cities and towns, a large general increase in literacy, as well as a shift in the content that readers were exposed to. Reading was no longer limited to canonical works such as the Bible; people increasingly read secular texts and novels, and they began to read for leisure. This reading revolution (e.g., Lagerfeld, 1986) not only changed what people read but how they read: alone and in silence where, for the first time, they had intimate access to the lives, experiences and perspectives of distant, unencountered others.

By exposing readers to the lives of others, literature began to raise awareness of the maltreatment of certain social groups. For example, the novels of Charles Dickens drew attention to poverty in London and the inhumane conditions of Victorian workhouses, and *Uncle Tom's Cabin* (Beecher Stowe, 1852/1995) highlighted the plight of slaves in the American South, helping to fuel the abolitionist movement of the 1850s, and lay the groundwork for the American Civil War in the 1860s (Kaufman, 2006). The humanitarian reform of the 1800s, which included improvements in women's rights, workers' conditions, and more humane treatment of criminals, coincided with the rise in novel-reading that followed the advent of mechanical printing (e.g., Pinker, 2011). This association suggests that

5Access to reading material varied across political climates. For example, by the end of the 18th century China was producing more books than other countries, but education was restricted to the elite and so the non-elite did not have the option to read until the 20th century (Lagerfeld, 1986).

either the growth of fictional stories as a form of entertainment was symptomatic of an increasingly socially conscious society, or the sudden proliferation of fiction-reading changed people's thinking on a mass scale, catalysing social reform.

> Stepping into someone else's vantage point reminds you that the other fellow has a first-person, present-tense, ongoing stream of consciousness that is very much like your own […] Slipping even for a moment into the perspective of someone who is turning black in a pillory or desperately pushing burning faggots away from her body or convulsing under the two-hundredth stroke of the lash may give a person second thoughts as to whether these cruelties should ever be visited upon anyone. (Pinker, 2011, p. 211)

According to Pinker (2011), by providing new perspectives through which to view the world, fiction-reading helped people to imagine themselves in the positions of others, thus reducing the appeal of, for example, violent acts like corporal punishment. Keen (2007) suggested that fiction enables the exploration of ideas and circumstances without the real-world consequences and obligations that may arise from reading nonfiction narratives. As Morgan (1963) put it: "a book is the only place in which you can examine a fragile thought without breaking it, or explore an explosive idea without fear it will go off in your face" (p. 4). Whereas nonfiction encourages the reader to connect the text to the real world and is more likely to induce action, fiction is more likely to induce empathy (Nünning, 2014), and provides a safe space wherein readers can exercise their "role-taking imagination" (Keen, 2007, p. 39). This process appears to begin at a young age, as an interest in fairytales typically starts to develop around age four, alongside the emergence of empathic abilities (Wimmer & Perner, 1983).

3.3 Fiction as Simulation of the Social World

Empathy has been associated with literary artistic practices since it originated as a concept. Plato predicted that a performer of literature who wept during a recitation would cause the audience to cry too. Aristotle observed that spectators of theatre would share in the

fear and suffering of characters due to their recognition that the acted scenes could reflect

real-world situations (Pinotti & Salgaro, 2019). Paraphrasing Robert Louis Stevenson, Oatley

(2011a) submitted that, "scientists use circles to solve problems in physics, and writers and

readers likewise use fictional characters to think about people in the social world" (para 1).

However, the cognitive processes involved in the effects of fictional works on empathy are, as

yet, unidentified.

Mar (2018a) proposed two routes through which stories may improve empathic

abilities: (i) via the presentation of social content leading to the accumulation of social

knowledge, and (ii) through the development of social processes, such as reasoning about

mental states, perspective-taking and emotional engagement. These accounts are analogous to

theory-theory versus simulation accounts of "theory of mind" or empathic inferencing (e.g.,

Coll et al., 2017; Goldman, 2006). According to the theory-theory approach, people develop a

theory about the way others think and behave. Through observation, children and adults

accumulate knowledge about the social world that enables those theories to be refined. This

generally accepted perspective was challenged when Heal (1986) and Gordon (1996) argued

that people understanding other minds should be thought of not as "theorisers" but as

"simulators" (though some debate was happening much earlier; see Stone & Davies, 1996).

The simulation theory proposes that people understand others by activating their own mental

processes—those processes that they would use in the situation at hand—and using them

online as proxies for those of the target (e.g., Coplan, 2011; Goldman, 1995). The discovery

of mirror systems in the motor region of the brain (Rizzolatti, Fadiga, Gallese & Fogassi,

1996) showed that perceiving an action was enough to elicit a mental simulation of that action

(Botvinick, Jha, Fabian, Solomon & Prkachin, 2005; Keysers, Wicker, Gazzola, Anton,

Fogassi & Gallese, 2004; see also Wicker, Keysers, Plailly, Royet, Gallese & Rizzolatti,

2003). This represented the first neural evidence in support of Lipps' theory that perceiving

and internally imitating an expression or behaviour would enable access to the emotion that caused it or the feeling of doing it (De Vignemont & Singer, 2006).

Goldman's (2006) account does not reject the theory-theory idea outright, rather it presents a hybrid view that understanding others involves both theorising and simulation processes, although the simulation mechanism is at the forefront of his proposition. This integrative view is also reflected in Mar's (2018a) model of fiction effects: the social content of fiction (i) provides knowledge that can be used to refine theories of how others think and behave and (ii) recruits the same mental apparatus as is involved in interpreting real-world social situations. Through these routes, fiction may support real-world empathic accuracy, and the two pathways may overlap: readers may develop social cognitive processes as well as a store of social narratives that they can subsequently draw on to understand a person's situation and experience (Gallagher & Gallagher, 2019). However, Mar (2018a) noted that most evidence in the fiction effects literature to date has supported the "process" account.

In line with the "process" account of fiction-engagement, Gerrig (1993) hypothesised that the same cognitive components are recruited when people process fictional events as when they encounter those events in the real world. Neuroscientific evidence suggests that cognitive processing of fictional narratives and of the experienced social world overlap: the same neural mechanisms are involved in processing real events and fictional descriptions of similar events (e.g., Speer et al., 2009; Wallentin, Nielsen, Vuust, Dohn, Roepstorff & Lund, 2011; see also Zwaan, 2004; for a meta-analysis, see Mar, 2011), and this effect is stronger for narratives than unconnected sentences (Kurby & Zacks, 2013). Oatley (2012) suggested that fictional narratives constitute valid representations of the experienced social world, and that because fiction is inherently social, it activates responses associated with real-world social interactions. In other words, fiction enables readers to simulate social experiences (Oatley, 1999).

This perspective is supported by evidence that brain areas (the anterior and mid-cingulate cortex) associated with experience-sharing (having an affective response similar to that of a target; Walter, 2012) are activated during immersion in fiction (Hsu, Conrad & Jacobs, 2014) and continue to remain active afterwards (Cadwell, 2015). Tamir, Bricker, Dodell-Feder and Mitchell (2016) found that reading and social cognition both recruit the brain's default network: a set of interacting brain regions associated with the ability to simulate hypothetical mental states and scenarios. Participants read passages that varied on imagery (vivid versus abstract) and content (mental content versus no mental content), while in an fMRI machine. Results showed that these dimensions led to patterns of brain activity differentially related to the simulation of vivid scenes versus the simulation of people and their minds. Moreover, their results indicated that the simulation of social scenes (identified via activity in the subnetwork associated with passages containing mental content) mediated the relationship between lifetime exposure to fiction and the tendency to spontaneously think about intentions when judging a target's behaviour in a vignette task. These findings suggest that the social simulacrum works at the neural level: in terms of processing, fictional and real events appear broadly indistinct.

Oatley and colleagues suggested that through simulating the social experiences depicted in fiction, readers may develop the skills associated with interpreting, making sense of, and responding appropriately to, social phenomena: "a play or novel runs on the minds of the audience or reader as a computer simulation runs on a computer" (Oatley, 1999, p. 105). In this view, fiction constitutes learning about the social world through the abstraction and simulation of social happenings, a form of "learning through experience", which can facilitate empathic growth (Mar & Oatley, 2008, p. 173). This, Mar and Oatley (2008) argued, represents the core function of fiction.

A growing field of research has examined the hypothesised positive relationship between fiction and empathic abilities. This chapter will now turn to a review of that literature. First, correlational studies examining relationships between lifetime fiction-exposure, fiction-engagement and empathic abilities will be reviewed, followed by experimental research examining causal effects of fiction-engagement on empathic abilities.[6] Evidence of other factors involved in the relationship between fiction and empathic abilities will also be addressed, particularly the role of immersion.

3.4 Review of Research Examining Relationships between Lifetime Fiction-exposure and Empathic Abilities

3.4.1 Bookworms versus nerds

If fiction is a simulation of the social world, then readers should show relative expertise in social interactions, just as a pilot-in-training can develop her skills in a flight simulator (Mar & Oatley, 2008). However, if a person spent all of their free time reading, time for interacting with people in the real world would be depleted. Mar and colleagues argued that such reading habits would not, necessarily, be detrimental to a person's social development (Mar & Oatley, 2008; Mar et al., 2006). Mar et al. (2006) conducted a study to test the stereotype of the socially awkward "bookworm" who has "chosen the company of print over his peers" (p. 695). They suggested that while this stereotype may have some basis in truth, it might only apply to avid readers of expository nonfiction, rather than to readers of novels. They argued that fiction, with its multiple intentional agents and social scenarios, may function as a substitute for real-world social interactions. On the other hand, expository texts, which do not imitate social environments, would not buffer against a lack of social

---

[6] Several literature searches were conducted. The latest systematic Boolean title search was conducted on 6th September 2019 on the Web of Science database using the following combinations of terms to search for research articles: fiction* or narrative or reading or story* or stories* or litera* and empath* or social cognit* or theory of mind or mentaliz* or mentalis* or prosocial or altruis* or social skills or social ability* or social acumen. This yielded 205 results, reduced to 120 following a title sift.

interaction. If this is the case, it would only be the nonfiction-reading "nerds" that would show reduced levels of social acumen, and not the fiction-reading "bookworms". Drawing on Oatley's (1999) simulation theory, they suggested that through fiction-reading, "bookworms" can simulate social experiences and develop their empathic abilities without needing to leave their armchairs.

To investigate this hypothesis, Mar et al. (2006) developed a new version of the Author Recognition Test paradigm (ART; Stanovich & West, 1989). ARTs provide a measure of lifetime exposure to fiction by quantifying familiarity with fiction authors as a proxy measure of lifetime exposure to fiction-in-print. The test consists of a list of names comprised of fiction authors and made-up, plausible foils. Participants are asked to select from the list those names that they know to be fiction authors. By incorporating foils, the task uses a signal-detection approach to discourage guessing and control for socially desirable responding (such as selecting unfamiliar names in order to appear more widely read). In the case of correctly selected names it is assumed that participants have read the author's books or become familiar with them when browsing related books in libraries, bookshops, online, and through adverts or recommendations. Although the task is an indirect measure of lifetime fiction-exposure, it has been shown to predict real-world reading abilities and habits (Mol & Bus, 2011; Rain & Mar, 2014; Stanovich & Cunningham, 1992; Stanovich & West, 1989; Stanovich, West & Harrison, 1995; West, Stanovich & Mitchell, 1993). The test has been revised to include updated author lists (as younger participants tend to recognise fewer of the original authors; Acheson, Wells & MacDonald, 2008) and genre dimensions. Mar et al.'s (2006) revision of the ART (ART-R) included both fiction ($n = 50$) and nonfiction ($n = 50$) author names, as well as plausible foils ($n = 40$). Both the fiction and nonfiction categories were divided into five genres to ensure that they represented the potential breadth of a person's reading habits

(fiction: romance, sci-fi/fantasy, suspense/thrillers, domestic fiction, foreign translated; nonfiction: science, philosophy/psychology, political/social commentary, self-help, business).

Mar et al. (2006) examined relationships between participants' fiction and nonfiction scores on the ART-R and self-report empathy (using the IRI) along with behavioural measures of mentalising (using the RMET and the IPT-15; see Chapter 2). The results showed that fiction-exposure was positively related to scores on the RMET, whereas nonfiction-exposure was negatively related to RMET scores. To return to the authors' original distinction, "bookworms" showed a higher ability to mentalise based on emotion expressions whereas "nerds" showed a lower ability to do so. While fiction-exposure did not relate to higher IPT-15 scores (the association was negative but non-significant) nonfiction-exposure was significantly negatively associated with scores on the task. These results held when controlling for the other dimension of the ART-R (fiction or nonfiction), foil-checking, age, years of English fluency and general intelligence. On this mentalising task, then, while fiction-exposure was unrelated to task performance, nonfiction-reading "nerds" performed worse.

Mar et al. (2006) found that associations between fiction-exposure (as well as nonfiction-exposure) and the self-report empathic ability scales were non-significant, except for fiction-exposure with the fantasy subscale of the IRI which, as described in the previous chapter, provides an index of the trait tendency to become immersed in stories. This suggests that people who become immersed in stories are likely to read more of them. With regard to the null findings for the other scales, the sample (94 participants) may have lacked sufficient power to detect effects. Indeed, a post-hoc power analysis conducted as part of the current literature review (using G*power; Faul, Erdfelder, Lang & Buchner, 2007) based on the largest non-significant raw correlation (fiction-exposure with the empathic concern subscale, $r = .09$)

indicated that power was low at .13.[7] Mar et al.'s (2006) study was the first to incorporate

fiction and nonfiction ART dimensions, and to use the ART to examine associations with

empathic abilities. Since then, studies investigating correlations with fiction-exposure have

also tended to use adaptations of the ART. These have facilitated the assessment of exposure

to different genres (Black et al., 2018; Fong et al., 2013, Kidd & Castano, 2017a) and have

been used to control for general fiction-exposure in studies examining the immediate, causal

effects of reading stories (Kidd & Castano, 2013; Koopman, 2016).

3.4.2 Meta-analysis of correlational studies (Mumper & Gerrig, 2017)

Mumper and Gerrig (2017) conducted a meta-analysis of published and unpublished

studies that used adaptations of the ART to examine relationships between leisure-reading and

empathic abilities. They included original research with healthy adult populations that had

measured associations between fiction or nonfiction-reading habits with self-report empathic

abilities ("dispositional empathy") and task-based empathic accuracy ("theory of mind"). The

studies included in the dispositional empathy analysis predominantly used the IRI (seven used

the empathic concern subscale, three used the fantasy scale, one used the personal distress

scale, and six used the perspective-taking scale). The behavioural empathic accuracy measure

was the RMET in all but one study that used a social vignette test (created by Young,

Bechara, Tranel, Damasio, Hauser & Damasio, 2010). In the latter, participants were asked to

view 48 vignettes and make judgments about the permissibility of an actor's behaviour. The

scenarios in the vignettes were varied by actor intention (negative, neutral) and outcome

(negative, neutral), and so participants' judgments about moral permissibility were considered

to reflect the extent to which participants spontaneously thought about the actor's intentions.

---

[7] The power of a statistical test is the probability that a false null hypothesis will be rejected. Power can be articulated as a value between 0 and 1 where higher values indicate higher power. Typically, power of $> .80$ is considered desirable in the social sciences, which sets the chance of wrongly accepting the null hypothesis at 20% or lower (e.g., Bakker, Hartgerink, Wicherts & van der Maas, 2016).

Therefore, the behavioural measures assessed either the explicit ability to decode facial expressions, or the spontaneous tendency to mentalise about an agent's intentions.

Mumper and Gerrig (2017) provided aggregate correlations between ART scores and the combined empathic ability measures (computed using the averages of the IRI empathic concern and perspective-taking scales or scores on the Toronto Empathy Questionnaire (TEQ; Spreng, McKinnon, Mar & Levine, 2009) and the aggregate empathic accuracy scores (RMET and social vignette task described above). ART scores were positively associated with dispositional empathy, $r(16) = .070$, $p < .001$, 95% CI [.03, .11], and with task-based empathic accuracy, $r(12) = .211$, 95% CI, $p < .001$ [.15, .27], and the association was stronger for the latter.[8] However, the measures used in the studies contributing to each of the aggregate effects tapped into the different empathic components (as outlined in Chapter 2). This heterogeneity is visible at the level of individual studies. For example, in their study examining associations between fiction-exposure and empathic abilities described in the previous section, Mar et al. (2006) observed no correlations between personal distress and any of the other IRI subscales. There was also no association between the behavioural tasks (the RMET and IPT-15), nor did the behavioural tasks correlate with the self-report IRI scales, except for a positive association between the IPT-15 and fantasy (which could represent a spurious correlation symptomatic of the small sample). This lack of correlation among the dependent variables highlights the heterogeneity of empathic ability measures (Mar et al., 2006) and the importance of a multidimensional approach.

Closer inspection of the individual self-report empathic ability scales illustrates this point. As identified in the previous chapter, the perspective-taking scale measures the tendency to mentalise about other people's points of view (e.g., "I try to look at everybody's

---

[8] The degrees of freedom were calculated as N-1 for each variable (N-2 in the bivariate correlations) where N is the number of effect sizes computed.

side of a disagreement before I make a decision"), whereas the empathic concern scale

measures feelings of sympathy (e.g., "I often have tender, concerned feelings for people less

fortunate than me"). The fantasy scale assesses the tendency to identify with fictional people

and situations, with all but one item ("I daydream and fantasise, with some regularity, about

things that might happen to me") pertaining to immersion in narrative worlds and empathy for

fictional characters (e.g., "after seeing a play or movie, I have felt as though I were one of the

characters"), and so the fantasy scale may reflect an experience-sharing process (although this

is debated; Jordan et al., 2016). The TEQ consists of 16 questions across the domains of

experience-sharing (e.g., "when someone else is feeling excited, I tend to get excited too"),

concern (e.g., "I have tender, concerned feelings for people less fortunate than me") and

mentalising ("I can tell when others are sad even when they do not say anything").[9] The

empathy subscale of the Literary Response Questionnaire (LRQ; Miall & Kuiken, 1995) was

used in four of the studies included in the meta-analysis. The scale contains seven items all

related to becoming absorbed in fictional stories with the majority of questions probing the

tendency to share in the experiences of characters (e.g., "when I read fiction I often think

about myself as one of the people in the story"). Therefore, this scale is most closely related to

the fantasy dimension of the IRI (although this association has yet to be tested) and may

reflect experience-sharing.

   Therefore, while aggregate correlations are useful for comparing the effect sizes of

fiction and nonfiction, the correlations between fiction-exposure and the individual empathic

ability scales would provide a more nuanced insight into the effects of fiction. Additional

correlations were reported for the IRI subscales, but not for the TEQ, LRQ, or the two

behavioural measures (the RMET and social vignettes test) individually. The IRI results

[9] This question could also reflect the sharing of emotion. However, the TEQ was found to correlate with
perspective-taking as well as the empathic concern subscale of the IRI (though the correlation was stronger with
empathic concern) indicating that some of the TEQ questions probe variance associated with the cognitive
component rather than with the emotional component exclusively.

revealed that fiction-exposure was positively associated with empathic concern, $r(16) = .074$, $p < .001$, 95% CI [.04, .11], and fantasy, $r(14) = .178$, $p < .001$, 95% CI [.14, .21], as well as perspective-taking $r(15) = .079$, $p < .001$, 95% CI [.04, .13], but not with personal distress $r(8) = -.042$, $p = .222$, 95% CI [-.11, .03]. This is unsurprising considering the difference between personal distress and the other constructs: the personal distress scale measures self-oriented responses to others' needs—i.e., the desire to reduce one's own distress—which can lead to avoidance of empathic behaviours, rather than other-oriented responses, distinguishing it from the other IRI dimensions (Baron-Cohen & Wheelwright, 2004; Batson & Shaw, 1991; Davis, Mitchell, Hall, Lothert, Snapp & Meyer, 1999; Penner et al., 1995; Singer & Lamm, 2009).

The positive relationships between reading and the individual empathic ability subscales that Mumper and Gerrig (2017) identified were not unique to fiction-exposure. Nonfiction-exposure showed smaller positive correlations with perspective-taking, $r(12) = .064$, $p = .003$, 95% CI [.02, .11], and fantasy, $r(10) = .054$, $p = .020$, 95% CI [.01, .10], and a significant negative correlation with personal distress, $r(8) = -.095$, $p = .011$, 95% CI [-.17, -.02]. There was no association between nonfiction-exposure and empathic concern, $r(12) = .035$, $p = .146$, 95% CI [-.01, -.08]. The aggregate correlations with task-based measures of mentalising (the RMET, $n = 14$, and social vignettes, $n = 1$) were stronger for fiction-exposure, $r(12) = .211$, $p < .001$, 95% CI [.15, .27], compared to nonfiction-exposure, $r(5) = .091$, $p = .008$, 95% CI [.02, .16]. Taken together, these findings indicate that fiction appears to be associated with processes involved in both the mentalising and experience-sharing domains. However, nonfiction also appeared to relate to mentalising abilities (RMET, social vignette task and perspective-taking scale) and less strongly to fantasy but not to empathic concern. Nonfiction was, however, negatively associated with the tendency to feel self-oriented personal distress. This suggests that some of the variance in mentalising and experience-sharing (if the fantasy

subscale is taken as an indicator of experience-sharing tendency) is accounted for by nonfiction-reading. However, the activation of the concern component at least may be unique to fiction. It could be that nonfiction leisure-reading cultivates the ability to reason about others' mental states and to imagine other worlds, whereas fiction may particularly invoke other-oriented affective or motivational processes.

3.4.3 Other correlational evidence from studies with adults

Mumper and Gerrig's (2017) meta-analysis provided a comprehensive review of correlational studies examining relationships between lifetime fiction- and nonfiction-exposure and empathic abilities. However, some experimental studies examining fiction effects (which will be reviewed in the next part of this chapter) have also included the ART, usually to control for variance associated with previous fiction habits, and they have also reported correlations with empathic ability measures. Kidd and Castano (2013; Experiments 1, 4 and 5) found that ART scores (Acheson et al., 2008) positively predicted RMET scores, $p < .01$, but not (Experiment 2) the Diagnostic Analysis of Nonverbal Accuracy 2-Adult Faces test (DANVA-AF; Nowicki, 2010, which measures facial emotion recognition using full faces). The positive association with RMET scores was also found in a replication of Kidd and Castano (2013) by Panero et al. (2016), $z = 4.46$, $p < .001$, and in a conceptual replication by Kidd, Ongis and Castano (2016), where higher scores on the ART led to higher RMET performance, $\beta = .25$, $p = .01$, and $\beta = .30$, $p < .01$, in two analyses using different stimulus texts.

Other studies have reported correlations between fiction and empathic abilities without using the ART. Johnson (2012) examined the relationship between empathic feelings experienced during reading a passage of fiction and performance on empathic ability measures. After reading, participants viewed morphed faces showing various levels of fear and happiness. Participants that experienced more concern-related feelings during fiction-

reading (e.g., "compassionate"; Batson et al., 1987) tended to say that fear was present on faces even when it was not: for each unit increase in emotions during reading, there was an 8% increase in participants' ratings for fearfulness, $p = .01$. In other words, they showed bias towards fearful facial expressions, thus accuracy for detecting fearful faces was lower for those that experienced emotions during reading (dropping by about 10% for each unit increase in emotions), $p = .016$. Accuracy for happy faces was not significantly affected. This is interesting for two reasons: first, it indicates that fiction-induced empathic processes that are not contingent on dispositional empathy (since trait concern was included in the regression models) relate to people's ability to mentalise about others' emotional states. Second, it suggests that the effects of empathic feelings during reading are not always positive. In this study, effects appeared to be influenced by context as they impacted the reading of fearful but not happy expressions; the bias was shown only for expressions that conveyed "need" (Johnson, 2012). As discussed in the previous chapter, concern is evoked by an awareness of need and, therefore, fiction-induced concern may predispose people to identify need from faces.

According to the empathy model established in the previous chapter, concern may lead to prosocial behaviour. Indeed, Johnson (2012) found that higher levels of fiction-induced feelings were associated with a higher likelihood of helping in the pen drop task outlined in the previous chapter (van Baaren et al., 2004): for each unit increase in emotions during reading, participants were 13% more likely to pick up the pen, $p = .042$). However, dispositional empathic concern was not associated with the prosocial behaviour outcome (although it was associated with the bias for fearful facial expressions). Therefore, some of the increase in prosocial behaviour may have been accounted for not by an increase in concern, but by participants modelling prosocial behaviour portrayed by the story characters

(Johnson, 2012). Further research is required to clarify the relationship between concern, fiction-induced feelings and prosocial behaviour.

3.4.4 Correlational evidence from studies with children

The focus of the studies described has been on NT adults which, for the reasons outlined in the previous chapter, is the target population of the present research. However, it is worth noting that correlational research has shown similar effects with children. Mar et al. (2010) examined 4-6-year-olds' levels of exposure to children's literature, television and film using four exposure measures based on the ART (authors and titles of literature, television programme titles and film titles), which were given to the participants' parents to complete. Five mentalising abilities were tested (the accurate attribution of desires, beliefs, knowledge, false beliefs and emotions). Additionally, parents completed the ART-R (Mar et al., 2006), which contained fiction and nonfiction authors of books for adults, as a measure of their own exposure to fiction. Controlling for vocabulary, age, gender and income, exposure to children's literature (assessed via parents' familiarity with children's literature) positively predicted children's mentalising scores, $\beta = .30$, $p < .05$, and this effect was found regardless of the parents' own exposure to fiction. Using the same control variables, a positive effect was also found for exposure to children's films, $\beta = .30$, $p < .05$, but not to children's television. As yet, it is unclear whether the same pattern is seen in adults.

Research has examined the effects of children's engagement with other fiction formats. In quasi-experimental studies with children and adolescents, Goldstein and colleagues examined the social benefits of acting compared to other types of arts classes. Goldstein et al. (2009) compared high school students taking elective acting classes with students from the same population not taking acting classes. They found a large effect of class taken on RMET performance, $d = 0.83$, $p = .008$, with acting students scoring higher on average. They also scored higher on a self-report empathy task (the Index of Empathy for Adolescents; Bryant,

1982) which measures empathic concern and experience-sharing through questions such as, "I think it is funny when I see someone cry" (reverse-scored), and "even when I don't know why someone is laughing, I laugh too". Goldstein et al. (2009) internally replicated this finding using the MASC (a film measure described in the previous chapter) and the empathic concern subscale of the IRI. Again, actors scored higher for mentalising, but this did not reach statistical significance, $d = 0.58$, $p = .057$. There was no difference between empathic concern scores for the acting and non-acting students (in fact, the acting students scored slightly lower on average).

In another quasi-experimental design, Goldstein and Winner (2012) compared children aged 6-11 who were already enrolled on or were recruited to acting classes versus visual arts classes. Before taking the classes, there were no differences between the groups on mentalising (measured using the Strange Stories Test, the RMET and the Faux Pas test which assesses the ability to recognise insults; Baron-Cohen, O'Riordan, Stone, Jones & Plaisted, 1999) or empathic concern (measured using the Index of Empathy for Children; Bryant 1982). However, the children in the acting course group did show higher levels of experience-sharing at baseline, measured using the Fiction Emotion-Matching task created for the study. In the task, participants were asked about the feelings of characters in film clips and their own emotional responses to them. This measure was scored according to matches between the emotions articulated by the participants for the targets and for themselves. After 10 months of participation, children on the acting course had made significant gains in their Index of Empathy scores (concern and experience-sharing) compared to the controls. On the Fiction Emotion-Matching task the actors scored significantly higher than the controls, $d = 1.23$, $p = .008$, but there was no interaction with time; in other words, the children enrolled on the acting class were more likely at both pre-test and post-test to share in the experiences of fictional characters presented on film. It may be, therefore, that children who tend to share in

the emotions of fictional characters are simply more drawn to taking acting classes, as the majority of participants self-selected into groups rather than being randomly assigned.

Children on the acting course did not show significant improvement compared to children on the visual arts course on any of the mentalising measures. The authors suggested that this was due to low ecological validity as the simplistic stimuli did not incorporate the range of visual and auditory cues that would be true to life. They conducted a follow-up study with secondary school students enrolled on competitive acting training, visual arts or music training (arts and music students formed the combined control group). They found that actors showed increases in mentalising, assessed using the Empathic Accuracy Paradigm described in the previous chapter, $d = 0.84$, $p = .026$, and a non-significant trend towards gains in trait experience-sharing, $d = 0.56$, $p = .06$, measured using the Basic Empathy Scale for Adolescents (Jolliffe & Farrington, 2006), which contains questions like, "after being with a friend who is sad about something, I usually feel sad". Again, the actors showed higher levels of emotion sharing on the Fiction Emotion-Matching task (this time using television programmes), $d = 0.73$, $p = .036$, though these differences were also found before training commenced, and they were not impacted by time. Actors also performed higher on the RMET before training. The findings from Goldstein and colleagues' studies show that children who signed up for acting courses tended to have higher baseline levels of experience-sharing, and adolescents who enrolled on acting training tended to show higher baseline levels of both mentalising and experience-sharing (Goldstein et al., 2009, Goldstein & Winner, 2012). Acting training appears to lead to gains in some empathic abilities above other arts training, specifically concern and experience-sharing in younger children, and mentalising and experience-sharing in adolescents. However, as the studies by Goldstein and colleagues were quasi-experiments with pre-determined groups (no random assignment), inferences cannot confidently be made about the direction of cause, and other factors such as social-

developmental opportunities and experiences may have come into play during the study durations.

Gentile et al. (2009) examined the effects of videogaming on the prosocial behaviour component, specifically whether video games with prosocial content were related to prosocial behaviours. In their study with Singaporean children (Study 1), participants' exposure to prosocial and violent video games was measured by asking them to list their three favourite games, the number of hours they spent on each weekly, and to rate how often players help or kill others in the game. This approach had previously demonstrated good construct and predictive validity (e.g., Anderson & Dill, 2000). Additionally, the authors checked validity by correlating the participants' ratings with expert ratings for the game, and agreement was satisfactory, $r = .75$, $p < .001$. Prosocial behaviour was measured using eleven helping items (e.g., "I would spend time and money to help those in need") and seven cooperation items (e.g., "I feel happy when I share my things with others") from the Prosocial Behaviour Questionnaire (Cheung, Ma & Shek, 1998), plus items adapted from the Children's Empathic Attitudes Questionnaire (Funk, Fox, Chan & Curtiss, 2008) related to trait empathy (e.g., "when I see a student who is upset, it really bothers me"), and other scales related to aggression and emotional awareness. Prosocial game playing was positively related to all helping behaviours, cooperation and sharing, and empathy. In their longitudinal study, Gentile et al. (2009, Study 2) asked Japanese children to rate how often in the past month they had played video games with prosocial scenes using 5-point Likert scales anchored with "not at all" and "very often". Here, prosocial behaviour was assessed via self-report: participants were asked how frequently they had performed prosocial acts (e.g., "I helped a person who was in trouble"). The authors identified a bidirectional relationship between prosocial gaming and prosocial behaviour, showing that prosocial game-play influenced behaviour 3-4 months later, $\beta = .069$, $p < .001$, and vice versa, $\beta = .074$, $p < .001$ (Study 1).

Recall that Johnson (2012) proposed that prosocial content might increase prosocial behaviour through a general learning mechanism. As it is not clear whether prosocial gaming increased empathy in Gentile et al.'s (2009) study, it may be that general learning accounts for some of the variance in prosocial behaviour that they identified. Taken together, correlational results from studies with children provide support for the hypothesis that fiction-engagement is positively associated with empathic abilities and prosocial behaviour and indicate that effects are not limited to reading. The extent that gains in empathic abilities are involved in fiction effects on prosocial behaviour, and the potential influence of other factors, remains unclear.

In conclusion, evidence from correlational studies supports the view that fiction and empathic abilities are associated (namely mentalising, experience sharing and concern), but the findings do not establish causation. It could be that reading fiction cultivates empathy, or that people high in empathic abilities tend to engage with more fictional stories (and empathic children with films; Mar et al., 2010). Alternatively, other factors may account for the relationship. For example, people with certain personality traits may be drawn to reading fiction, and those same traits may make them more empathic. However, fiction-exposure is not likely to be self-selected by young children. While children with more developed empathic faculties could request access to certain fiction media or drama classes, these requests are not necessarily met by caregivers (Mar et al., 2010). Although this provides some indication of a causal effect of fiction, findings from studies with children do not necessarily generalise to adults. Several researchers have used experimental approaches with NT adult groups in order to test the hypothesis that reading passages of fiction causally influences empathic abilities. This chapter will now turn to a review of that literature.

3.5 Review of Research Examining Causal Effects of Fiction-engagement on Empathic Abilities

3.5.1 Literary fiction improves "theory of mind"

A causal link between fiction and mentalising ability was established by Kidd and Castano (2013). They hypothesised that reading literary (acclaimed or canonical) fiction, would causally impact empathic accuracy. Across five experiments, participants in the "literary" condition (Experiments 1-5) were randomly assigned to read one of a set of short stories or passages from novels by contemporary award-winning authors (PEN/O Henry Award for short literary fiction or US National Book Award for Fiction winners) or acclaimed authors (e.g., Anton Chekhov). In the "popular" condition (Experiments 2-5) participants read passages from Amazon bestsellers (e.g., *Gone Girl;* Flynn, 2012) or from stories featured in an anthology of popular fiction (Hoppenstand, 1998). In the nonfiction condition (Experiment 1) they read factual narratives (e.g., *How the Potato Changed the World*, Mann, 2011, or *The Story of the Most Common Bird in the World*, Dunn, 2012). After reading, participants took part in a series of empathic accuracy tasks, a false belief task, the RMET, DANVA-AF and Yoni test. In Experiments 2 and 5, participants in a control condition did not engage in any reading task.

Results showed that RMET scores were higher for people that read literary fiction compared to nonfiction (Experiment 1, $\omega_{p2} = 0.05$, $p = .01$), popular fiction (Experiments 3 and 4, both $\omega_{p2} = 0.04$, both ps $= .04$), and either popular or nonfiction (Experiment 5, $\omega_{p2} = 0.01$, $p = .04$). The authors suggested that reading literary fiction had recruited the social cognitive mechanism associated with interpreting others' mental states and thereby improved performance on the mentalising task post-reading. The findings gained a great deal of traction and the study has been cited well over 1000 times. However, two initial replication studies (combined $N = 1798$) failed to reproduce the finding that reading literary fiction immediately

enhances mentalising, finding no effect of literary fiction compared to nonfiction, popular fiction and no reading (Panero et al., 2016; Samur et al., 2018).

There are several reasons that replication attempts may fail. Effect sizes in the original study were small, and so low statistical power may have been a factor (Samur et al., 2018), although the replications counteracted this problem via high-powered studies. Kidd and Castano (2017b) argued that the replication efforts did not sufficiently imitate the original study methods. Specifically, they suggested that Panero et al. (2016) did not ensure participants read their assigned texts, that participants who did not complete the ART were not excluded and, most importantly, that two of the replication experiments did not satisfactorily use random assignment, indicated via highly unbalanced groups (the nonfiction condition contained around three times the number of participants in the literary and popular conditions in each experiment). Despite these legitimate concerns, Panero et al. (2017) noted that Kidd and Castano's (2017b) re-analysis of their data still did not yield the same pattern of results as the 2013 study. Instead, RMET performance was higher for literary fiction compared to popular fiction (LSM_{10} = 27.21 and 25.60 respectively, p = .04), but not compared to the expository nonfiction-reading control group: participants in the nonfiction group outperformed those in the popular fiction group (LSM = 28.32 and 25.60 respectively, p = .004; though this latter effect was moderated by fiction-exposure). Furthermore, scores in the nonfiction group were no different compared to the literary fiction group. This suggests that there was an advantage to reading nonfiction compared to popular fiction, and no disadvantage compared to literary fiction.

It could be that expository and literary texts contain similarly complex linguistic devices responsible for effects on mentalising and that there is no benefit to reading literary fiction over literary nonfiction. Kidd and Castano (2017b) described this as an "unexpected finding"

---

[10] Least squares mean: the mean estimate adjusted for covariates and missing data in the model.

(p. 4) and referred to Kidd and Castano's (2013) finding that literary fiction and nonfiction did differ on the RMET (Experiment 1), as support for the claim that reading literary fiction benefits mentalising relative to nonfiction. However, a subsequent replication of Experiment 1 was conducted by Camerer et al. (2018) as part of a series of replication attempts of twenty-one experiments published in the prestigious journals *Science* and *Nature* between 2010 and 2015, and this failed. In response, Kidd and Castano (2018a) argued that Experiment 1 was the least refined of their (Kidd & Castano, 2013) original series of five discrete experiments and not the most appropriate to replicate. In light of this, the nonfiction comparisons and their implications warrant further investigation.

Kidd and Castano (2018a) argued that their (Kidd & Castano, 2013) fifth experiment had been refined as a result of the previous four sets of findings and so it was higher-powered, the hypothesis was narrowed, and it entailed more rigorous checks (e.g., for inattentive participation). They suggested that the fifth experiment would therefore be more reliably replicated than the first. Kidd and Castano (2018b) conducted three pre-registered replications of the fifth experiment. They used an analytical approach (Small Telescopes approach; Simonsohn, 2005) to compare the effects they identified to those which would have been detected using only 33% power in the original study (based on a sensitivity analysis of the original effects). This identifies whether a replication effect serves as evidence for the null hypothesis, or whether it is simply inconclusive. Using this technique, their third study was a successful replication and the other two were characterised as inconclusive.

Publicity surrounding the Kidd and Castano's (2013) study focused on the interesting, statistically significant findings for the RMET, but discounted the null findings for the DANVA-AF and Yoni test. Further, in Experiment 5, the inclusion of covariates took the significance value for the effect of fiction condition on RMET scores over the alpha threshold, $p = .06$, indicating that the effect was unstable. Van Kujik, Verkoeijen, Dijkstra and Zwaan

(2018) conducted a *p*-curve analysis of Kidd and Castano's (2013) original results. The *p*-curve analysis procedure was developed by Simonsohn, Nelson and Simmons (2014) to examine the distribution of significant *p*-values in order to identify whether a set of significant findings is the result of *p*-hacking or selective reporting. It indicated that the evidential value of their findings was low. One explanation noted by van Kujik et al. is that Kidd and Castano included different covariates in their analysis of the RMET data than they did for the Yoni test data, which was unexplained. Nevertheless, their replication (they replicated Kidd and Castano's, 2013, Experiment 5, to follow up the *p*-curve analysis) did identify a significant effect of condition on both the RMET and Yoni test cognitive content scores, in line with Kidd and Castano's original pattern of results. Kidd et al. (2016) also conducted a successful conceptual replication of Kidd and Castano's (2013) finding that reading literary fiction (using a different text selection) leads to higher RMET scores (*LSM* = 26.22) compared to popular fiction (*LSM* = 24.71), *p* = .02, and to no reading (*LSM* = 24.64), *p* = .05.

Evidence from the replication attempts is certainly mixed, including several null results. The prospect that the original positive findings could have represented Type 1 errors raises doubt not only about the effect of literary fiction, which will be returned to later in this chapter, but about the impact of fiction in general.

3.5.2 Meta-analysis of experimental studies (Dodell-Feder & Tamir, 2018)

Dodell-Feder and Tamir (2018) conducted a meta-analysis of the effects of fiction-reading on empathic abilities. They included only true experiments (where participants were randomly assigned to conditions) that measured empathic abilities which, in line with Mumper and Gerrig's (2017) meta-analysis of correlational research, included tests of the ability to perceive and interpret social information and tests of responses to that information (based on the definition by Fiske & Taylor, 2013). Their meta-analysis included 14 studies which yielded 53 individual effect sizes derived from fiction-reading (*n* = 1615) and

nonfiction-reading or no reading control participants ($n = 1843$). The range of empathic ability measures used were categorised according to Zaki and Ochsner's (2012) model differentiating mentalising and experience-sharing. The mentalising category included behavioural (e.g., RMET) and self-report measures (e.g., IRI perspective-taking subscale). The experience-sharing category included self-report measures only (e.g., the Multifaceted Empathy Test-Emotional Empathy component; Dziobek et al., 2008, which quantifies the strength of affective response to images of individuals in distress).

The results of the meta-analysis supported the primary hypothesis that participants that read fiction would score higher on the empathic ability measures compared to participants that read nonfiction or did not read, Hedge's $g = .15$, $p = .029$, 95% CI [.02, .09]. A more granular inspection of the results facilitates a nuanced assessment of the relationship between fiction-reading, mentalising and experience-sharing processes, prosocial concern and behaviour. Table 1 shows the effect sizes included in the meta-analysis arranged by empathic ability (assessed post-reading), method (behavioural or self-report) and the precise measures used to assess each. The presentation of results here diverges from Dodell-Feder and Tamir (2018) in order to reflect the model of empathy established in the previous chapter. Specifically, the empathic concern measures have been separated from the experience-sharing measures in order to represent the separate concern component. Average effect sizes have been calculated for mentalising, experience-sharing, prosocial concern and prosocial behaviour and, then, for measurement format (self-report or behavioural) where applicable. One result was omitted from these calculations as it was not clear which questionnaire items were used in the study (the EQ–Emotional Reactivity scale used by Pino & Mazza, 2016, $g = 0.14$).

The average effect size for comparisons between fiction and nonfiction or no reading was $g = 0.60$ ($SD = 0.58$), for experience-sharing. One particularly large effect size may have positively biased the result ($g = 1.51$). Excluding this value, the average was 0.37 ($SD = $

0.33). For mentalising measures, the average effect size was 0.12 (SD = 0.32). Effects were larger for behavioural (M = 0.14, SD = 0.34) compared to self-report tasks (M = -0.02, SD = 0.21), for which the average effect was negative, indicating that the scores for participants in the fiction conditions were lower than for the comparison groups. There were more positive effects than negative effects for self-report tasks, but larger effects on the negative side. This indicates that of the studies included in the meta-analysis, behavioural tests of mentalising were positively impacted by fiction-reading, although when reflecting on their abilities, participants did not necessarily consider themselves to be more skilled or more frequent mentalisers. All but one of the self-report mentalising measures used the perspective-taking subscale of the IRI. As this provides an index of the dispositional tendency to see things from others' points of view, it may not detect immediate fiction effects.

The average effect size for the concern measures (which were all based on fiction versus nonfiction comparisons) was also negative (M = -0.03, SD = 0.32). Only two of the six effect sizes were negative, but one particularly so, g = -0.55. Again, most of the measures used were dispositional concern scales (and both negative values were measured using the IRI empathic concern scale). Fiction-reading had a positive effect on prosocial behaviour (g = 0.40) although only one effect size was provided. Fiction also positively impacted experience-sharing on average (M = 0.60, SD = 0.58) and no values were negative (the excluded study value was also positive). This aggregation of effect sizes across empathy components reveals that the largest effect of fiction across studies was on experience-sharing. The positive effect of fiction on mentalising was only present when behavioural measures were employed. Self-report measures were useful for assessing correlations between dispositional empathic abilities and fiction-exposure but may not be sensitive to immediate change caused by fiction manipulations.

Table 1

Measures and Effect Sizes Adapted from Dodell-Feder and Tamir's (2018) Meta-analysis. Effect Sizes are Organised First by Empathic

Component, then Measure Format, Measure, Effect Size, Study Details and Comparison Group.

Empathic ability	Method	Measure	Effect size (Hedges' *g*)	Author(s), year.	Comparison group	Mean effect sizes (*SD*)
Mentalising						Mentalising: *M* = 0.12 (0.32)
	Behavioural	Faces Test (Identify emotion from picture of actor's facial expression)	0.95	Pino & Mazza, 2016	Nonfiction	Behavioural: *M* = 0.14 (0.34)
		Social-Reasoning Task (Infer mental state of character in story)	0.66	Mar, 2007	Nonfiction	
		Reading the Mind in the Eyes Test (RMET: recognise emotions from eye regions of faces)	0.54	Kidd & Castano, 2013	Nonfiction	
		RMET	0.37	Samur et al.	No reading	
		RMET	0.36	Kidd et al., 2016	No reading	
		RMET	0.34	Samur et al.	Nonfiction	
		RMET	0.28	Kidd & Castano, 2013	No reading	
		RMET	0.22	Samur et al.	No reading	
		RMET	0.17	Samur et al.	Nonfiction	
		RMET	0.16	Black & Barnes, 2015b	Nonfiction	
		RMET	0.08	Djikic et al., 2013	Nonfiction	
		RMET	0.06	Samur et al.	Nonfiction	
		RMET	0.03	Panero et al., 2016	No reading	
		RMET	-0.04	Panero et al., 2016	Nonfiction	
		RMET	-0.07	Weisberg	No reading	
		RMET	-0.08	Panero et al., 2016	Nonfiction	
		RMET	-0.10	Liu & Want	Nonfiction	
		RMET	-0.15	Panero et al., 2016	Nonfiction	
		RMET	-0.23	Samur et al.	Nonfiction	

	Diagnostic Analysis of Nonverbal Accuracy – Affective (DANVA-AF: recognise emotions from whole faces)	0.44	Kidd & Castano, 2013	No reading
	DANVA-AF	0.33	Panero et al., 2016	No reading
	Multifaceted Empathy Test – Cognitive Empathy (MET-CS: identify mental state from pictures)	0.24	Pino & Mazza, 2016	Nonfiction
	Emotion Attribution Task (identify character's emotion based on vignette of situation)	0.16	Pino & Mazza, 2016	Nonfiction
	Imposing Memory Task (IMT: infer beliefs and emotions of story characters; increasing complexity)	0.10	Panero et al., 2016	No reading
	Yoni test (infer thoughts and emotions to schematic face based on eye gaze and facial expression	0.05	Kidd & Castano, 2013	No reading
	Yoni test	-0.22	Panero et al., 2016	No reading
	IMT	-0.02	Weisberg	No reading
	IMT	-0.24	Panero et al., 2016	Nonfiction
	False belief task (infer character's belief or action based on their false belief)	1.05	Pino & Mazza, 2016	Nonfiction
	False belief task	-0.13	Kidd & Castano, 2013	No reading
	False belief task	-0.20	Kidd & Castano, 2013	Nonfiction
	False belief task	-0.50	Panero et al., 2016	Nonfiction
Self-report				Self-report: M = -0.02 (0.21)
	Interpersonal Reactivity Index – Perspective-taking (IRI-PT: report tendency to take others' perspectives)	0.17	Djikic et al., 2013	Nonfiction
	IRI-PT	0.17	Weisberg	No reading
	IRI-PT	0.06	Mar	Nonfiction
	IRI-PT	-0.13	Johnson et al., 2013	Nonfiction
	IRI-PT	-0.22	Johnson et al., 2013	Nonfiction
	IRI-PT	-0.33	Liu & Want	Nonfiction
	Empathy Quotient – Cognitive Empathy (EQ-CE: report ability to take others' perspectives)	0.16	Pino & Mazza, 2016	Nonfiction

Experience- sharing						Experience-sharing: $M = 0.60$ (0.58)
	Self-report	Empathy (rate understanding for a group's experience [marginalised social group or depressed/bereaved individuals])	1.51	Johnson et al., 2013	Nonfiction	
		Empathy	0.83	Johnson et al., 2013	Nonfiction	
		Empathy	0.06	Koopman, 2015	Nonfiction	
		Multifaceted Empathy Test -Implicit Emotional Empathy (MET-IEE: indicate emotional arousal in response to images of individuals in distress)	0.36	Pino & Mazza, 2016	Nonfiction	
		Multifaceted Empathy Test -Explicit Emotional Empathy (MET-EEE: indicate strength of feeling for image of individuals in distress)	0.23	Pino & Mazza, 2016	Nonfiction	
		Empathy Quotient – Emotional Reactivity (EQ-ER)	0.14	Pino & Mazza, 2016	Nonfiction	Not included in calculation
Prosocial concern						Prosocial concern: $M = $ -0.03 (0.32)
	Self-report	IRI-Empathic Concern (IRI-EC: report concern or sympathy for others)	0.25	Mar	Nonfiction	
		IRI-EC	0.25	Liu & Want	Nonfiction	
		IRI-EC	0.14	Djikic et al., 2013	Nonfiction	
		IRI-EC	-0.28	Bal & Veltkamp, 2013	Nonfiction	
		IRI-EC	-0.55	Bal & Veltkamp, 2013	Nonfiction	
		Toronto Empathy Questionnaire (TEQ: empathic responses to others)	0.00	Koopman, 2015	Nonfiction	
Prosocial behaviour						Prosocial behaviour: $M = 0.4$ (0.0)
	Behavioural	Charity donation (opportunity to give participation compensation to charity)	0.40	Koopman, 2015	Nonfiction	

Note. Effect sizes (Hedge's g) represent the standardised mean difference between fiction-reading and comparison group in post-reading empathy measures. Positive values indicate better performance in the fiction-reading group. Mean effects sizes and standard deviations are presented for each component, and by measure format within components. Authors without dates represent unpublished studies.

3.5.3 Experimental findings from studies with children

Studies with children have provided some evidence of effects of engaging with fiction on social abilities (e.g., Goldstein et al.'s, 2009, and Goldstein & Winner's, 2012, quasi-experiments described earlier). Larsen, Lee and Ganea (2017) tested the impact of reading about human characters versus anthropomorphised animal characters on children's prosocial behaviour. In a control condition, children read a story about seeds. Afterwards, all of the children were invited to choose ten stickers as a reward for taking part. The experimenter explained to each participant that another (anonymous) child would receive no stickers as they had not taken part, and the participant was then given the opportunity to secretly share any number of the stickers that they had chosen with that child. The number of stickers given away constituted the prosocial behaviour measure.

Children who read about human characters were more likely to share their stickers after reading ($M = 2.91$, $SD = 2.37$) compared with children who read about animal characters ($M = 1.69$, $SD = 1.73$) or seeds ($M = 1.53$, $SD = 1.98$). P-values were calculated by comparing the mean differences in sticker-sharing in each condition and were $p = .001$ for the human compared to animal condition and $p = .002$ for the human compared to control (seeds) condition. The authors concluded that realistic stories were important for promoting prosocial behaviour. As there was no explicit measure of empathy, it is unclear how far it would factor in this effect. There were instances of prosocial behaviour depicted in the stories, and so it could be that the children simply replicated the prosocial actions of the characters. If this were the case, realism may have been important to the children's tendencies to identify with the characters and their social context, allowing them to accumulate applicable social knowledge, and this may have mediated the effect on prosociality. Future research would be required to test empathy and identification with characters as potential contributors to the effect on behaviour.

Schellenberg (2004) assigned 6-year old children to either drama lessons, music lessons or a waiting list control group for a 36-week period. The music group was the critical condition in this study, and children in this group showed an increase in IQ compared to the other two groups. However, children that received drama lessons showed an improvement in social abilities (measured using a composite score of parent ratings of a range of social behaviours) from before to after the intervention, $p < .001$, that was not evident in the music groups, and a comparatively large improvement compared to the other groups, $d = .057$, $p < .01$. This could have been due to the narrative content engaged with via drama, which included characters and their mental states. Alternatively, it could have been the result of other aspects of the drama class, such as improvisation, engaging with imagined props, or interactions with other children. Finally, it may have resulted from external factors that were unrelated to the classes. All classes contained the same number of children and so the results appeared specific to the drama context rather than to the size of the group, although the processes responsible for the observed effects remain opaque. Studies with children suggest that fiction may impact social skills and behaviours, though they have not provided answers to the question, arising from both the correlational and causal literature, of which factors are involved in such effects.

3.6 Factors Involved in Fiction Effects

3.6.1 Publication status and individual differences

Mumper and Gerrig's (2017) meta-analytic procedure revealed small but significant relationships between fiction and empathy. However, a limitation of meta-analyses is that they inevitably reflect the quality of the data entered into the analysis (Esterhuizen & Thabane, 2016). This can result in positively biased results (larger than the true effects) due to an over-representation of significant findings in the literature since null findings are often unpublished ("file-drawer effect"; Rosenthal, 1979). Mumper and Gerrig (2017) included nine

unpublished studies in their analysis, which were obtained through communications with authors of published findings. Publication status had no effect on fiction habits or dispositional empathy, although the authors acknowledged that there may be more unpublished studies on the topic that were not located. In Dodell-Feder and Tamir's (2018) meta-analysis of causal effects, published studies yielded larger effect sizes compared to unpublished studies, though publication status did not moderate the effect of fiction on empathic accuracy. Both meta-analyses (Dodell-Feder & Tamir, 2018; Mumper & Gerrig, 2017) found no evidence of a file-drawer effect.

In the causal meta-analysis, Dodell-Feder and Tamir (2018) noted significant heterogeneity among effect sizes which they attributed to within-study variance caused by the different formats of the empathic ability measures, or by different comparison groups. For this reason, the authors examined the impact of six moderators: publication status (published, unpublished), sample (online via Mechanical Turk, student sample), comparison group (nonfiction, no reading), format of empathic ability measure (self-report, behavioural), empathic ability (mentalising, experience-sharing), and participant characteristics (age, gender). None of these factors moderated the small effect of fiction on empathic ability aggregate scores. Gender and study design did not moderate the correlations in Mumper and Gerrig's (2017) meta-analysis either, lending evidence to a systematic relationship.[11]

Individual differences that were not addressed in the meta-analyses may also impact fiction effects on empathic abilities. Nettle (2006) showed that professional actors, whose careers involve engaging with fictional narratives via play-texts and performing as characters, tend to score higher on the Empathizing Quotient (EQ; Baron-Cohen, Richler, Bisarya,

---

[11] It is notable that empathic domain (mentalising or experience-sharing) also did not appear to moderate the effect; however, there was some overlap within the empathic ability categories, with concern and experience-sharing grouped together. The approach taken in this chapter distinguished empathic concern from experience-sharing and looked at average correlations for different measure formats within these classifications (whereas Dodell-Feder & Tamir, 2018, compared measure formats across the whole sample).

Gurunathan & Wheelwright, 2003) compared to general population samples, $d_{males} = .049$, $p <$ .001; $d_{females} = .057$, $p < .001$. This finding could reflect a third variable, a personality trait that is associated both with empathic abilities and with being drawn to the acting profession. Similarly, the children that signed up for drama classes in Goldstein and colleagues' studies described above (Goldstein et al., 2009, Goldstein & Winner, 2012) showed higher levels of empathic abilities compared to controls both before and after training. Therefore, a dispositional factor may both increase the tendency to choose drama classes and improve performance on empathic ability tasks.

Mar, Oatley and Peterson (2009) attempted to rule out these kinds of individual differences. They examined relationships between fiction-exposure and the Big Five personality traits (Costa & McCrae, 1992). Their pilot studies showed that fiction-exposure was related to trait openness to experience, $r = .22$, $p < .05$, but not to the other four traits (conscientiousness, extraversion, agreeableness and neuroticism). People high in this trait tend to show higher levels of open-mindedness, motivation to seek new experiences, creativity and absorption in activities (particularly artistic activities). Controlling for this trait and other demographic differences (age, gender, English fluency and IRI fantasy scores), fiction-exposure still predicted performance on the RMET, such that for each author recognised RMET scores increased by 7%, $p < .05$.

Mar et al.'s (2009) findings, coupled with those of Dodell-Feder and Tamir's (2018) and Mumper and Gerrig's (2017) meta-analyses, indicate that demographic differences, Big-Five personality traits and the publication status of results do not account for relationships between fiction-exposure and empathy dimensions.

3.6.2 Content and format

3.6.2.1 Fiction format: medium. Although fiction- and nonfiction-exposure tend to correlate (people that tend to read more, read more of both; Mar et al., 2006, 2009) the results

of Mumper and Gerrig's (2017) meta-analysis showed that lifetime exposure to fiction uniquely accounted for some variance in empathic abilities above and beyond exposure to nonfiction. This indicates that fiction specifically, rather than general reading processes, is related to empathy.

Mar et al.'s (2010) research with children showed that correlations differed depending on story format (exposure to books and film were positively associated with empathic abilities but exposure to television was not). In adults, causal research has indicated that the effects of fiction do not appear limited to prose; for example, viewing theatre has been shown alter people's attitudes and behaviours (e.g., Heide, Porter & Saito, 2012; Stephens-Hernandez et al., 2007; Valente & Bharath, 1999). Black and Barnes (2015a) used the same paradigm as Kidd and Castano (2013) to examine the effects of television dramas. Results showed that participants performed better on the RMET after watching award-winning TV dramas compared to documentaries, $d = .043$, one-tailed $p = .017$. However, the version of the ART used (Acheson, et al., 2008) and entered into the statistical model as a covariate assesses familiarity with fiction authors and not with TV dramas. As familiarity with TV dramas was not captured by the measure and therefore not controlled for, causation is less certain. It is unclear whether there would be effects of watching fiction on other empathic ability outcomes.

The aphorism, "the medium is the message" (McLuhan, 1964/1994), encapsulates the notion that media presentation is fundamental to the way content is received.[12] Different fiction media likely incur different levels of engagement (e.g., van Laer, De Ruyter, Visconti & Wetzels, 2014): viewing media, for example, may not require full attention, which can facilitate other cognitive processes (e.g., memory production; McDonald, Sarge, Lin, Collier,

---

[12] For the title of a subsequent publication (McLuhan, Fiore & Agel, 1967/2008), McLuhan revised the phrase to "the medium is the *massage*" (my italics). According to the McLuhan estate, the latter title was an error in typesetting that McLuhan liked, and kept, as the play on words captures media's effects on the whole sensorium (McLuhan, n.d.).

& Potocki, 2015; see also Mar et al., 2010). A comprehensive account of fiction effects must address the multifaceted nature of fiction-engagement.

3.6.2.2 Fiction content: genre. Kidd and Castano (2013) suggested that literary fiction utilises devices such as unusual linguistic features, inconsistent characterisation and gaps in the narrative that provide "grist for the mills" (Zunshine, 2006, p. 16) of empathic abilities (see also Hakemulder, 2000; Koopman & Hakemulder, 2015). Drawing on Barthes' (1974) distinction between "readerly" and "writerly" texts, this process, they argued, requires readers to take a creative, "writerly" approach to actively constructing the inner lives of complex characters, thereby recruiting the cognitive mechanisms associated with attributing mental states to others. In contrast, popular fiction is "readerly", containing consistent and predictable characters, so the need to interpret their mental states in order to make sense of the narrative is reduced. Along the same lines, Storr (2019) described the distinction between mass and literary media in terms of how moments of change are approached in the storyline: "change in mass-market story is quick and clear and easily understandable, while in high literature it's often slow and ambiguous and depends on plenty of work from the reader, who has to ponder and de-code the connections for themselves" (p. 54). He highlighted that this distinction is not limited to novels, but also applies to film media. Art-house films such as the works of David Lynch, for example, may disrupt the logic of cause and effect, requiring the viewer to decode the chronology of events, characters intentions and experiences.

Historically, fiction effects research lacks systematic comparisons of literary and non-literary fiction. Studies have tended to involve variables other than literariness (Koopman & Hakemulder, 2015), resulting in mixed evidence for the importance of literary features. Kidd and Castano (2017a) conducted a factor analysis—an approach to identifying the underlying factor structure of a measure—on the ART (Acheson et al.'s, 2008, version) in order to distinguish familiarity with literary versus non-literary genre-fiction (see also Moore &

Gordon, 2015). Controlling for age, gender and dispositional empathic concern (IRI subscale), familiarity with literary fiction authors positively predicted mentalising (RMET) scores, $\beta = .22$, $p = .018$, but familiarity with genre-fiction authors did not. De Mulder, Hakemulder, van den Berghe, Klassen and Van Berkum (2017) found that literary fiction-exposure correlated with mentalising about cognitive content on the Yoni test described in the previous chapter, $r = .21$ (p-values were not provided), but not affective content (Yoni test affective component or RMET). Because the design was correlational, the direction of cause could not be determined.

Kidd and Castano's (2013) experiments, described earlier in this chapter, showed a positive causal effect of fiction on mentalising across five discrete experiments. In addition to the failed replications addressed above (section 3.5.1), there are two reasons to question the specificity of the effect to literary fiction: first, the fiction-exposure measure (Acheson et al., 2008) did not distinguish between literary and popular fiction, and so prior exposure to the different genres was not controlled for. Second, the contents of the text stimuli were not matched across conditions, and so the observed effects may not have been specific to literariness but, rather, artefacts of engaging with particular themes (Koopman, 2016).

This raises the problem of classifying literariness. Keen (2011) criticised distinctions between "high and "low" literature but identified that literary fiction contains "complicated, changeable, and category-resistant characters" in contrast to the "stereotyped characters engaged in predictable actions" of popular fiction (p. 303), or complex "round" versus simplistic "flat" characters, to use E.M. Forster's (1927/1974) distinction. However, she also suggested that it is precisely because simpler characters' experiences and struggles are unambiguous, that they may invoke greater empathy than more complex characters. Kidd and Castano's (2013) stimuli were fiction from "literary" and "popular" anthologies. Establishing the extent to which the stories selected actually differed in terms of content would require a

linguistic or thematic analysis. A range of texts were selected for each group, which would have reduced the impact of potentially confounding differences between single stimulus texts, although the statistical procedure did not account for this source of heterogeneity within groups (Panero et al., 2016).

An alternative approach to that taken by Kidd and Castano (2013) is to examine particular textual features or themes, rather than classifications of literariness. There are several aspects of style which can be used to distinguish literary from popular fiction (Kidd et al., 2016). Koopman (2016) investigated the effects of literary features by providing one group of participants with a text containing "foregrounding". Foregrounding refers to striking literary features that depart from everyday language use and are subjectively experienced as making a text seem more original (Koopman, 2016; for examples of text with and without foregrounding, see Kuzmičová, Mangen, Støle & Begnum, 2017). A second group of participants was assigned to read versions of the same text without foregrounding (in which common alternatives were used in place of semantic, phonetic and grammatical foregrounding) and a third without imagery (in which metaphors were replaced by literal alternatives). Results showed that participants that read the original story with foregrounding, compared to those that read the version without foregrounding, scored higher on self-reported understanding of the experiences of people in similar situations to the protagonists, $d = 0.86$, $p < .01$. For this measure, participants indicated agreement with statements on 7-point Likert-style scales. Statements included inferencing items (e.g., "I can understand people who, multiple years after a beloved has died, are still preoccupied with the loss") and items indicating prosocial concern (e.g., "the basic insurance policy should cover people who keep struggling with loss after many years"). Internal consistency was somewhat low, $a = .60$, though effects were still significantly lower for the version without foregrounding. This indicated that uniquely literary features may be responsible for some increase in

comprehension of, and concern for, people's experiences after reading about characters in similar situations. It remains unclear whether foregrounding specifically, rather than other literary features, accounts for such gains in empathic abilities.

Along similar lines, Hakemulder (2004) found that a text about immigration that contained foregrounding led to more positive attitudes towards immigrants compared to a version of the text in which foregrounding had been reduced. Participants read versions of passages from Rushdie's (1988) *The Satanic Verses*, which features two Indian immigrants being tracked by police until one of them is wrongfully arrested. Foregrounding was reduced in one condition via the deletion of descriptive detail and unusual metaphor. This shortened the sentences, resulting in a simpler summary of events. A third (control) group of participants did not read either text. Participants were subsequently asked to indicate agreement with items about attitudes towards immigrants in Western Europe. Readers of the original text (with foregrounding) indicated beliefs that intolerance towards migrants was more problematic compared to readers of the text without foregrounding and participants that did not read, $d = 0.48$ and 0.90 respectively, $p < .05$, assuming equal group sizes. However, this effect was not observed for three other scales about attitudes towards migrants: their motives in relocating, difficulties faced in their new environments, and levels of acceptance of acceptance from their new communities. The latter showed a significant effect of reading group but differences disappeared when planned contrasts were conducted. This could have been due to low power which led to a weakly significant initial omnibus test, $p < .05$ (exact *p*-value unavailable). In light of this, the extent that the story altered readers' schemas regarding immigration is not clear.

Kuzmičová et al. (2017) showed a different pattern of results. They gave participants either a literary text or a non-literary version of a text in which foregrounding was reduced. Participants were first asked to mark passages that they found striking or evocative in the text,

and then asked to elaborate on three of the sections that they had marked. The version *lower* in foregrounding led to more explicitly empathic elaborations compared to the original story. Van Peer, Hakemulder and Zyngier (2007) found no impact of foregrounding on their affect measures (feeling "moved", "touched" and "sad"), though Koopman (2016) pointed out that this may be due to the affect items measured, or due to only one line of text being manipulated across conditions. Evidence that foregrounding is the aspect of literary fiction that impacts empathy is, so far, inconclusive.

In a conceptual replication of Kidd and Castano (2013), Kidd et al. (2016) found that participants who read literary texts scored higher on the RMET (*LSM* = 26.22) compared to those who read popular fiction and (*LSM* = 24.71), p = .02, and a baseline (no reading) condition (*LSM* = 24.98), p = .05. Their supplementary linguistic analysis examined reflective function using a computer programme that quantifies mental state terms used reflectively, such as, "I felt really angry and then it changed to sadness" (as opposed to the non-reflective mere presence of mental state terms, e.g., "I was angry, now I am sad"). They analysed each of the texts used by Kidd and Castano (2013) to demonstrate that reflective function was higher in the literary condition compared to the popular condition, p = .044. Further, a correlation analysis positively related reflective function to RMET scores, r = .48, p = .02. This suggests that the inclusion of mental state terms may support mentalising when linked to reflective markers, and that these tend to appear in canonical texts and literary award winners (as this is how the texts were classified in the original experiment).

Djikic and Oatley (2014) argued that literary fiction provides an artistic experience that activates personality fluctuations, and that these fluctuations can lead to long-term change. Djikic, Oatley, Zoeterman and Peterson (2009) found that a short story by Chekhov prompted significantly greater self-reported personality change than a content-matched documentary version; however, it is unclear whether this change was contingent on literary features (such

as foregrounding) specifically, or narrative features more generally (the presentation as fiction). Personality was measured as an overall change in the Big Five Inventory (BFI; John, Donahue & Kentle, 1991), which is a 44-item self-report measure of the Big Five personality traits. The authors ran a one-tailed test (*t*-test) despite absolute scores on the BFI measures having been calculated due to the lack of a directional (positive versus negative) prediction, which increases the possibility of a Type 1 error. Even so, significance was only at trend level for the full, composite BFI, $p = .053$, and the findings were non-significant for each of the five individual scales. A regression analysis showed an association between condition and BFI change. Here, significance was at the alpha threshold, $p = .05$, and so further testing would be required to assess the stability of the effect. Emotion change was measured before and after reading and consisted of self-report questions about the intensity of ten emotions (sadness, anxiety, happiness, boredom, anger, fearfulness, contentment, excitement, unsettledness, and awe) at the time of completion. There was significant emotional change for participants that read the literary text rather than the documentary version, $p < .009$ [*sic*]. The predictive effect of condition on personality change (established via regression) did not reach significance when the variance caused by emotion change was controlled. Therefore, reading material appeared to lead to changes in the intensity of emotional states, but the effect on personality was less certain.

Black and Barnes's (2015a) finding that award-winning TV dramas positively influenced emotion recognition compared award-winning dramas with documentaries and not with other (non-acclaimed) narrative dramas. Therefore, the narrative element of the stimuli, rather than artistic features, may have been responsible for the observed effect. However, Djikic, Oatley and Moldoveanu (2013) found that perceiving a text as "artistic" and "interesting" predicted self-reported increases in perspective-taking ability measured using the IRI subscale, $r = .25$, $p < .05$, whereas whether a text was fiction or nonfiction was not a

significant predictor. As both artistry and perspective-taking were subjective measures, whether change actually occurred, and whether this change was prompted by literariness, remains speculative.

In a correlational study, Fong et al. (2013) examined adults' exposure to four genres: domestic fiction, romance, science-fiction/fantasy, and suspense/thriller, and their relationships with mentalising (RMET scores). When other factors (age, gender, English fluency, exposure to nonfiction, and personality measured using the BFI) were controlled, only the romance and suspense/thriller genres were associated with the ability to mentalise about emotions, $r = .17$, $p < .01$, and $r = .12$, $p < .03$, respectively, with domestic fiction at a trend level of significance, $r = .10$, $p < .08$. The authors noted that 95% confidence intervals for domestic fiction did not include 0, though they did for suspense/thriller. They suggested, therefore, that the correlations concerning these two genres were less certain compared to the association between exposure to romance stories and mentalising.

Examining other outcome variables, Fong, Mullin and Mar (2015) correlated the same four genres as Fong et al. (2013) with sexual attitudes, though they only observed associations for the composite fiction-exposure measure (which was associated with increased gender role egalitarianism and fewer gender role stereotypes). However, Black et al. (2018) measured exposure to seven genres and found that people who recognised more literary, science-fiction and fantasy authors indicated that they considered more morally dubious scenarios to be permissible, whereas people who recognised more romance authors found *fewer* scenarios morally permissible. It is unclear if, and how far, egalitarian attitudes and moral permissibility judgments are associated with empathic abilities, although these findings indicate that associations with fiction-exposure can vary, not only in terms of literary versus popular classifications, but at the level of thematic genres.

The importance of artistic, literary features and their associated processes in effects on empathic abilities is far from clear. On the one hand, studies have indicated that devices such as complex, unpredictable characters and striking aesthetic features in literary fiction are key to augmenting social abilities; on the other hand, there is evidence that non-literary fiction has positive effects on social skills. It remains possible that fiction-induced improvements in empathic abilities are entailed by features that are often associated with, but not exclusive to, literary texts (such as foregrounding). Addressing the issue of defining literary fiction, Koopman (2018) pointed out that it is not enough for a literary text to deviate from "everyday" language, it must also be noticed by the reader (here, she draws on Miall & Kuiken, 1994). Participants' awareness of the artistry of the text may contribute to its effects and, conversely, readers may be able to engage with popular fiction in "literary" ways: literary fiction may require deep engagement to produce effects on empathy, and popular fiction may not *require* this to be understood but may *allow* it (Barnes, 2018). The question to answer, then, is not whether literary fiction is more suited to creating change than popular fiction, but which specific features, themes or processes produce such effects.

3.6.3 A possible role for immersion

Dodell-Feder and Tamir (2018) pointed out that although their meta-analysis provided evidence of a positive, causal effect of fiction-reading on empathic abilities, it did not provide information about the processes involved in these effects, and they suggested that future research should aim to elucidate these. Most research evidence to date supports the role of fiction-engagement processes (rather than the influence of specific fictional content; Mar, 2018a) in fiction effects on empathy. However, further clarification of the nature of these processes and their effects on specific dimensions of empathy is needed. As discussed earlier in this chapter, it has been suggested that literary devices such as striking aesthetic features and complex, unpredictable characters require readers to fill in "gaps" in the narrative, and

entail an active, "writerly" approach to interpreting characters' inner experiences. This recruits and improves mentalising processes (De Mulder et al., 2017; Hakemulder, 2000; Kidd & Castano, 2013; Pino & Mazza, 2016; Zunshine, 2006). Such narrative devices are typically associated with literary fiction (acclaimed or canonical works; Gavaler & Johnson, 2017; Hakemulder, 2000; Kidd & Castano, 2013); however, studies reviewed here have also linked popular genre-fiction with empathic abilities (e.g., Fong et al., 2013), as well as prosocial behaviour (Koopman, 2015) and moral reasoning (Black et al., 2018). Moreover, suspenseful narrative, which can feature in popular thrillers, has been linked to activity in areas of the brain associated with social cognition (Lehne, Engel, Rohrmeier, Menninghaus, Jacobs & Koelsch, 2015). Taken together, these findings suggest that fiction effects on empathic abilities are not contingent on processes unique to literary reading.

A possible solution to this problem arises from the field of narrative persuasion research. Although the messages in fiction tend to be implicit, fictional stories can alter people's attitudes, beliefs and behaviours in predetermined directions (Slater, 2002). This can occur via analytic (Chaiken, 1987; Escalas, 2007; Petty & Cacioppo, 1986) and narrative persuasion processes (Green & Brock, 2000; Green, Garst & Brock, 2004). Analytic persuasion assumes the analysis of explicit messages, whereas narrative persuasion facilitates the impact of implicit messages through a non-critical, experiential process of becoming immersed or "transported" into a story (Gerrig, 1993). As discussed earlier, Gerrig's (1993) transportation theory was fundamental to Oatley's (1999) simulation account of fiction effects on empathic inferencing. However, despite the theory of fiction as simulation being quite well accepted, the role of immersion in fiction effects on empathy has not been sufficiently addressed.

Green and Brock (2000) posited transportation as the process through which narrative persuasion occurs. They developed a measure of transportation comprising eleven primary

questions about quality of engagement, such as, "when I was reading the narrative, I could easily picture the events in it taking place", and four additional questions relating to story-specific content, for example, "while reading the narrative, I had a vivid image of the boy" (Green & Brock, 2000, p. 704). The scale, therefore, predominantly measures mental imagery, and one question probes emotional engagement with characters (Busselle & Bilandzic, 2009; Green & Brock, 2002). Results showed that higher transportation led to more story-consistent beliefs (e.g., after reading a story about the murder of a young girl by a psychiatric patient, highly transported participants were more likely to agree with restrictions to psychiatric patient freedom). While attitudinal change indicates that fiction can have transformative effects, it does not necessarily signify a shift in empathic abilities. It has been found, however, that increasing empathy for individual members of marginalised groups improves attitudes towards the group as a whole (people with HIV/AIDS and people experiencing homelessness; Batson et al., 1997), and so empathic processes may be implicated in attitude change. Furthermore, effects could vary between readers: as relationships between fictional characters, their thoughts and feelings, are often implicit, fiction may generate change not through persuasion but through indirect communication. In other words, the direction of change may not be predetermined, but may vary in relation to the specific mind engaging with it (Djikic & Oatley, 2014).

If transportation can facilitate the persuasive impact of implicit messages via a peripheral, as opposed to a central, elaborative route (Petty & Cacioppo, 1981), non-critical engagement processes may also facilitate the development of empathic abilities. As Oatley's theory suggests, the more engaged people feel with story characters and events, the more likely they are to take those experiences with them to the world outside the story (Oatley, 1999, 2012; Mar & Oatley, 2008).

Immersion in stories is conceptually linked with empathic processes. As discussed in Chapter 2, the fantasy subscale of the IRI measure of empathy tends to be treated as a measure of trait transportability or imaginative abilities (e.g., Mar et al., 2006; see also Hall & Bracken, 2011). All but one item refers to immersion in stories and empathy for fictional characters, and scores positively predict transportation (e.g. Mazzocco, Green, Sasota & Jones, 2010). This trait has been shown to correlate with both self-report and behavioural measures of empathy (e.g., Mar et al., 2006, identified a positive correlation with performance on the Interpersonal Perception Task-15, $r = .28$, $p < .05$, as well as with the perspective-taking, $r = .24$, $p < .05$, and empathic concern, $r = .42$, $p < .05$, dimensions of the IRI). The inverse of this relationship has also been shown: Gavaler and Johnson (2017, Experiment 1) found that adding mental state explanations to narratives increased transportation, $d = 0.48$, $p = .004$. Thus, immersion appears to be associated with empathy-related content.

Bal and Veltkamp (2013) provided some evidence of a causal effect of immersion on empathic abilities. In Study 1, participants read a chapter from a fictional story by Arthur Conan Doyle, or two news stories from a newspaper. Results showed that participants reported higher concern a week after reading, though not immediately after reading. This was only true for fiction-readers and only if they had been transported into the story: the interaction effect was significant, $\beta = .17$, $p < .05$. In Study 2, participants read a chapter from a novel by José Saramago, or a set of five newspaper stories. Results showed that increases in transportation enhanced concern for fiction-readers, whereas nonfiction-readers higher in transportation showed a decrease in empathy, $\beta = .18$, $p < .05$. Non-transported fiction-readers in both studies showed lower empathic concern after a week. The impact of transportation in both studies was only qualified through the interaction with condition, and it is difficult to draw conclusions based on the fiction versus nonfiction comparison as the texts were not well-matched. For example, in Study 2, transportation was significantly higher in the fiction

condition, which is unsurprising given that the fiction condition contained a single narrative, whereas the nonfiction condition contained a selection of discrete stories (in Study 1, the two conditions were not compared for transportation). While change in empathic concern from time 1 (baseline) to time 3 (one week after reading) was graphically depicted, time was not included in either analysis as a within participants variable. Therefore, it is not clear how far individuals' empathic concern changed over time and the extent that this change was associated with transportation levels. While these problems are not inconsequential, the findings merit further research into the role of transportation in fiction effects on empathic concern, and on other empathic domains.

Kidd and Castano (2013) found that transportation during reading did not correlate with empathic accuracy. However, it is important to note that this was not the focus of their study, and so transportation was not experimentally manipulated (and did not differ significantly between fiction-reading conditions). Experiments 1 and 2 revealed no relationship between transportation and mentalising about cognitive content, but the measure, a false belief task, suffered ceiling effects with the NT adults tested. Mentalising about affective content was measured using the RMET in Experiments 1, 3, 4 and 5, and the DANVA-AF in Experiment 2 (effects of reading literary texts were significant for the RMET but not the DANVA-AF). The RMET and the DANVA-AF test both measure the ability to recognise mental states from the faces of actors (eye regions versus whole faces respectively). As empathy is multidimensional, it may be that immersion impacts some but not other empathy components. Experiment 5 identified a non-significant trend towards higher transportation in the literary ($M = 3.90$) compared to popular condition ($M = 3.83$ [SDs unavailable]) $p = .07$, though a comprehensive investigation of the effect of transportation on empathic accuracy would require transportation to be manipulated between groups.

Johnson's (2012) correlational study showing fiction effects on prosocial behaviour found that transported readers experienced more sympathetic emotions towards story characters (e.g., "compassionate"; Batson et al., 1987) and that this accounted for some of the increase in prosocial behaviour after reading. Johnson, Cushman, Borden & McCune (2013) followed this study with an experimental approach where they assigned participants to one of three groups. In the critical imagery-generation condition participants were given instructions on generating imagery and asked to do so while reading the story. In two control conditions, participants were asked to focus on verbal-semantic information (the meaning of particular words within sentences), or to read the text in the same way that they would usually read a story for leisure. Results showed that participants who generated more imagery while reading the narrative became more transported. They were over three times more likely than participants in the leisure-reading group to behave in a prosocial way by responding to a call for survey participation by an ostensibly unfunded researcher for a nominal fee of $0.05, $z = 2.47$, $p = .007$. The comparison with the verbal-semantic condition approached, but did not reach, significance, $z = 1.59$, $p = .056$. A limitation of the study was that it did not include a manipulation check (to assess how far participants had followed the instructions and either generated imagery or focused on the meaning of words). A failure of the manipulation could explain the lack of difference between the visual-semantic and leisure-reading groups.

Most important for the current argument, the authors conducted a mediation analysis which revealed that transportation fully mediated the effect of the imagery training on affective empathy for the story characters (using the same measure from Batson et al., 1987, as the Johnson, 2012, study described above). Both transportation and affective empathy mediated effects on prosocial behaviour ($z = 2.3$, $p = .01$, and $z = 2.0$, $p = .02$, respectively). This indicates that imagery when reading promotes immersion, which leads to prosocial behaviour both directly, and indirectly through increasing affective empathy for story

characters. The story was specifically designed to elicit empathy and included prosocial content. The authors suggested that their findings were consistent both with Oatley's (1999) simulation theory, and with a general learning approach (e.g., Gentile et al., 2009) in which the content of a story produces an internal state that leads to a behaviour. In this case, the internal state produced by the content of the story was empathic concern for the characters.

In contrast, Bormann and Greitemeyer (2015) found that immersion did not account for increased RMET scores caused by playing video games containing storytelling features. Participants either engaged with or ignored (via a distraction task) the narrative features of a video game which had received critical acclaim for its excellent narrative. In the game, the player is required to take on the role of a character whose quest is to resolve the mystery of her missing family. A third group of participants played a different game that involved climbing a virtual wall and did not involve other characters, explicitly or implicitly. As levels of social content were not comparable across the two games, Bormann and Greitemeyer called for future research to explore the mechanisms responsible for the impact of the narrative game. Furthermore, their immersion scale incorporated nine items which assessed physical (e.g., "when moving through the game world I feel as if I am actually there"), emotional (e.g., "I experience feelings as deeply in the game as I have in real life"), and narrative immersion (e.g., "when playing the game I feel as if I was part of the story"). The nine items were averaged to generate a composite immersion score, but the underlying constructs could impact empathy dimensions in different ways; engaging with the narrative could activate some but not other dimensions of immersion and these, in turn, could impact some but not other dimensions of empathy. In other words, the composite measure may fail to capture effects both on and of individual dimensions of immersion.

Narrative immersion is multifaceted: it involves character identification (reacting to and reacting with characters; Iguarta, 2010), mental imagery, and affective responses (e.g., Green

& Brock, 2002). These constructs have rarely been studied simultaneously. Differences in these processes may lead to different outcomes concerning empathy domains. Taking a multidimensional approach to both narrative immersion and empathic accuracy would help to establish how far dimensions of immersion impact different inferencing processes.

It is not only the field of research examining fiction effects on social cognition that would benefit from an understanding of the fiction-engagement processes that impact empathic abilities. As addressed in Chapter 2, several tasks commonly used to measure these skills employ narratives as stimuli. For example, the SST requires participants to read a fictional narrative about a couple whose romantic relationship breaks down (Hemingway, 1988), and empathic accuracy is assessed via a range of questions about the mental states of the story characters. In the original validation study Dodell-Feder et al. (2013) found positive correlations between trait fantasy, RMET scores and scores on the SST questions. If participants high in trait fantasy became more transported in the stimulus story, this could have impacted their empathic accuracy scores (this was not measured in the study, but fantasy has been shown to predict transportation; Mazzocco et al., 2010). Individual differences in levels of immersion may influence empathic accuracy and, if so, immersion should be accounted for in empathy tests which use narratives as stimuli.

3.7 Making Sense of the State of the Art

Getting "lost" in a good book (Nell, 1988, p. 8) provides a distraction from people's everyday lives but does not necessarily entail any psychological benefit. Fiction could represent nothing more than a "virus of the mind" that has hijacked humans' adaptive capacity to communicate complex social information (e.g., Brodie, 2009). However, research indicates that there may be benefits to engaging with fiction. Self-report and behavioural data have captured associations between exposure to fiction and both mentalising and experience-sharing abilities. Nonfiction-exposure, too, appears to positively relate to mentalising (though

not to empathic concern) but correlations are weaker than for fiction. Therefore, the evidence indicates that reading in general is associated with advantages in empathic abilities, but variance in some empathic abilities is uniquely attributable to fiction. Furthermore, fiction-reading may benefit some empathic components in particular. The breakdown of Dodell-Feder and Tamir's (2018) meta-analysis of causal effects (Table 1) revealed that the effects of fiction-reading appear to be largest for experience-sharing measures.

It is not all about reading, however; at least, it should not be. Fiction varies both in presentation format and in content. Research has reported positive associations in child populations between empathic abilities and engagement with fiction via books, film, and through acting. Experimental research has indicated that reading books, acting, and viewing some TV dramas, are linked to higher empathic abilities in adult populations. How far exposure to different modes of fiction relates to empathic abilities remains unclear. The question of which themes or genres may particularly support empathic abilities is also, as yet, unanswered, though research has indicated that effects are more nuanced than broad "literary" versus "popular" classifications or "high-brow" versus "low-brow" evaluative distinctions. Taking a more granular approach to measuring both fiction and empathic abilities will contribute to an understanding of how far the content presented in fiction, and the processes used to engage with that content (Mar, 2018a), account for fiction's effects.

Experimental studies have revealed small but positive effects of fiction on empathic abilities. A nuanced reading of meta-analytic data indicates that effects are largest for self-report experience-sharing, followed by mentalising measured via behavioural tasks, and that there are also positive effects on prosocial behaviour. This foregrounds the multidimensionality of empathy. The range of measures employed to assess empathy constructs has led to heterogeneity both within Dodell-Feder and Tamir's (2018) meta-analysis and across the literature. The current literature indicates that effects of fiction on

empathic abilities are generally positive but small. This does not mean that they are not important. As Dodell-Feder and Tamir pointed out, empathic abilities are associated with a range of positive outcomes including social connection (e.g., Goldstein, Vezich & Shapiro, 2014), wellbeing (e.g., Ryan & Deci, 2000) lower levels of illness (Yang, Boen, Gerken, Li, Schorpp & Harris, 2016) and even longevity (Holt-Lunstad, Smith & Layton, 2010). Considering the regularity with which people can access fictional narratives, as well as the extent that empathy is associated with positive outcomes, a small effect can be considered meaningful.

In spite of the potential social impact of fiction, the processes involved in its effects on empathy remain opaque. As the experiential process of becoming immersed in stories has been shown to increase narrative persuasion, and in light of some evidence that immersion may increase empathy, this mechanism warrants further investigation. Taking a multidimensional approach to both fiction and empathic abilities could provide insight into the mediators of their relationship. In turn, this would help to identify the contexts in which fiction may be most beneficial for empathic abilities.

3.8 Summary and Moving Forward

This chapter introduced the definition of fiction adopted in the current research. Correlational and causal studies investigating relationships between fiction and empathic abilities were reviewed. It was proposed that a range of factors may influence fiction effects, particularly media format, genre, and immersive processes, and that research should explicitly address the multidimensionality of each construct and approaches to their measurement. This may help to make sense of inconsistencies across the literature and to clarify avenues for further research. The next chapter presents the rationale and hypotheses for the three phases of empirical work conducted that aimed to address these issues.

Chapter 4: Rationale and Objectives

The previous two chapters established definitions of fiction and empathy, discussed traditional approaches to measurement, and used this information to frame a review of correlational and experimental studies that have investigated relationships between fiction-engagement and empathic abilities. It was stipulated that future research should address the multidimensionality of both in order to elucidate the processes involved in their associations. This chapter will present the rationale for the empirical research conducted in the current project. First, the multidimensional approach identified in the previous two chapters will be refreshed. Next, three studies will be defined, which aimed to clarify the nature of relationships between fiction and empathy. Main hypotheses will be determined, and further predictions and exploration outlined. Next, general methodological principles and potential implications of the research will be identified. A table of the full set of hypotheses will be provided at the end of the chapter. These can be tracked through the following three chapters, which function both as standalone empirical reports, and which document the progression from each study to the next.

4.1 The Multidimensional Approach

Fiction is multidimensional: it can be engaged with via a range of media formats, and story content varies across thematic genres. Empathic abilities, too, are heterogeneous and dissociable: skills in emotion recognition are not equal to skills in mentalising about a target's false beliefs, or the tendency to feel concern towards others in need. Empathy can be considered not as a unidimensional construct, but as an umbrella term encompassing the range of processes that support people's understanding of others (e.g., de Waal, 2011). The variety of fiction stimuli used in fiction effects research, and the range of empathy dimensions tested, may have contributed to heterogeneity across effects reported in the fiction-empathy literature, and rendered the processes involved in these effects opaque. The literature to date

has identified small, positive effects of fiction on empathic abilities (e.g., Dodell-Feder & Tamir, 2018), and so this line of enquiry need not be abandoned, rather it requires a nuanced approach. While all studies in the present project were designed to address the complexities of fiction and empathy, the first study focused on the multidimensionality of fiction, the second on the multidimensionality of empathy, and the third on the immediate effects of immersion in fiction, taking the multidimensionality of immersion into account. In so doing, the present research aimed to contribute to knowledge of how far fiction-engagement benefits empathic abilities, and the conditions under which effects occur.

4.1.1 Study One: Fiction is multidimensional

4.1.1.1 Fiction media-exposure and empathic abilities. Studies reviewed in Chapter 3 have shown that general exposure to fictional stories is positively associated with empathic abilities, indicating either that fiction inculcates these skills or that empathic people tend to read more fiction. Such studies have used variations of the ART paradigm (Stanovich & West, 1989) described in the previous chapter, in order to quantify exposure to printed fictional prose. However, people engage with fiction via a range of media formats: they may read stories, view them onscreen or live at the theatre, for example. Some studies have identified relationships between engagement with fiction via other media channels (e.g., television dramas; Black & Barnes, 2015a) and empathic abilities, though, as yet, there exists no tool to test adults' exposure to fiction via non-reading formats. The first study in the present doctoral research entailed the production of such a tool, designed to provide a composite measure of exposure to fictional stories via a range of media formats, as well as to enable the examination of individual relationships between fiction media dimensions and empathic abilities. Based on previous evidence that film and theatre can impact attitudes and empathy, it was hypothesised that, in addition to print-exposure, film and play-exposure would also positively predict empathic abilities. Inclusion of these dimensions would go some

way toward representing the range of formats through which people commonly engage with fiction. Mentalising, experience-sharing, concern and prosocial behaviour components, as identified in the empathy model developed in Chapter 2, were measured using the well-established IRI self-report scales, and the self-report altruism scale (all described in Chapter 2). In order to examine how far multidimensional fiction-exposure relates to the components of empathy, the following hypotheses were tested:

H_1: Fiction-exposure via print, film and plays combined would positively predict perspective-taking.

H_2: Fiction-exposure via print, film and plays combined would positively predict fantasy.

H_3: Fiction-exposure via print, film and plays combined would positively predict empathic concern.

H_4: Fiction-exposure via print, film and plays combined would positively predict altruism.

Additionally, individual associations between the fiction media-exposure scales and empathic abilities were explored to examine the strength and direction of associations across the dimensions of fiction-exposure and empathy. Furthermore, as participants may not have equal access to all of the media formats measured in the test of fiction media-exposure (e.g., due to proximity to a theatre), they were also asked about their preferences for different media formats. These preferences were examined for correlations with empathic abilities. This also enabled relationships between fiction-exposure and fiction preferences to be tested.

4.1.1.2 Enjoyment of fictional genres and empathic abilities. Medium may impact the effects of fictional social content on empathic abilities, though there remains debate about the nature of that social content. Chapter 3 reviewed the argument that literary fiction, due to its complex characters and unpredictable narratives, is uniquely positioned to recruit the cognitive mechanisms associated with understanding others (e.g., Hakemulder, 2000; Kidd & Castano, 2013). While there is evidence to support this claim (e.g., Kidd & Castano, 2013;

Kidd & Castano, 2017a; Djikic & Oatley, 2014; van Kujik et al., 2018), other studies have failed to reproduce these effects (Camerer et al., 2018; Panero et al., 2016; Samur et al., 2018; see also Panero et al., 2017; Kidd & Castano, 2017b; Kidd & Castano, 2018b). Chapter 3 identified potential reasons for inconsistencies in the fiction effects literature, including the limitations of cross-sectional research designs examining immediate effects with heterogeneous sets of short fiction extracts (fiction effects may require an incubation period, for example; Appel & Richter, 2007). Study 1 addressed this in two ways: first, by examining associations with long-term fiction-exposure (correlations have been shown in several studies) and, second, by examining a range of genres in order to reflect the themes and tropes that non-technical, broad "literary" versus "popular" classifications fail to distinguish. This approach was motivated by studies that have demonstrated relationships between social abilities and non-literary fiction genres (e.g., Black et al., 2018; Fong et al., 2013), as well as specific themes and features (e.g., depression and grief, Koopman, 2015; and suspenseful storylines, Lehne et al., 2015). Study 1 aimed to establish how far particular empathic abilities are associated with non-literary story types, in addition to examining how far fiction media formats predict empathic abilities. This was addressed through an exploration of relationships between people's preferences for different thematic genres and their empathic abilities.

4.1.1.3 Experience of acting and empathic abilities. In participatory engagement with fiction through acting, empathy can function not only as an *outcome* of engaging with fictional characters but as a *method* (Gallagher & Gallagher, 2019). Previous research, outlined in Chapter 3, has identified positive relationships between acting experience and empathic abilities. There is some evidence that enrolment on acting classes may support empathic development in children (Goldstein & Winner, 2012; Goldstein et al., 2009; Schellenberg, 2004), and professional actors tend to show higher self-report empathic abilities compared to the general population (Nettle, 2006). However, it is unclear how far experience

with acting relates to different empathic dimensions, and so the present study also aimed to explore relationships between acting experience and empathic abilities. The findings could lay the groundwork for future research aiming to compare fiction-engagement modes—reading novels, watching actors portray characters on film or at the theatre, and portraying characters first-hand—and their effects.

The first study aimed to scope out relationships between fiction media-exposure, media and genre preferences, acting experience and empathic abilities. Based on the empathy framework established in Chapter 2, it aimed to measure self-report mentalising, experience-sharing and concern (using the IRI which was featured in several of the studies reviewed in Chapter 3), as well as the prosocial behaviour component. To the author's knowledge, this represents the first study to examine relationships between exposure to fiction via different media and empathic abilities, and to include the prosocial dimension.

4.1.2 Study Two: Empathy is multidimensional

In Chapter 2 it was argued that empathy is not something one has or does not have; rather it encompasses a set of inferencing processes that can be enacted with different degrees of success. Empathic accuracy represents the measurable ability to accurately interpret a target's internal mental states using those inferencing processes (Ickes, 1997). The empathy model described in Chapter 2 illustrates that empathic accuracy can be achieved using mentalising or experience-sharing inferencing strategies. Once a target's mental state has been identified via one of these routes, prosocial concern and behaviour may be activated (Zaki & Ochsner, 2012). Recent research has indicated that the two strategies for empathic accuracy differ in value: Zhou et al. (2017) found that sharing in the experiences of targets led to more accurate mental state inferences compared to explicitly interpreting targets' facial expressions (mentalising about affective content) and that participants tended to undervalue the more successful experience-sharing approach.

4.1.2.1 Contextual features of stimuli. As explained in Chapter 2, Zhou et al. (2017) conceded that while "simulation" (the experience-sharing strategy) was shown to be the more effective strategy in their set of experiments, this may not be the case across all situations. Contextual features that can have an effect on empathic inferencing include characteristics of the experiencer (e.g., via an ingroup advantage; Adams et al., 2010; Matsumoto et al., 2009) and the valence of the target mental state content (Kauschke et al., 2019). Thus, the present study aimed to replicate Zhou et al.'s (2017) finding of an effect of strategy, while also accounting for differences in valence and target experiencer, as well as to assess any interactions. Specifically, the following hypotheses were tested:

H5: There would be an effect of empathic strategy on empathic accuracy, with participants that used the experience-sharing strategy showing higher empathic accuracy compared to participants that used the mentalising strategy.

H6: There would be an effect of valence on empathic accuracy.

H7: There would be an effect of experiencer on empathic accuracy.

H8: There would be an interaction effect of strategy, valence and experiencer on empathic accuracy.

4.1.2.2 Fiction-exposure and empathic accuracy. As yet, no studies have examined how far fiction-exposure relates to mentalising versus experience-sharing strategies via behavioural rather than self-report approaches. In Dodell-Feder and Tamir's (2018) meta-analysis of causal effects, reviewed in Chapter 3, mentalising was measured using a range of self-report and behavioural tasks, but all studies that measured experience-sharing did so using self-report scales. The second study aimed to address this gap in the literature. This line of enquiry reflects a novel mapping of Zhou et al.'s (2017) behavioural paradigm onto the components of empathy identified in the model outlined in Chapter 2.

If fiction is a simulation of the social world, then participants who share in the experiences of fictional characters may be more adept at using their own mental states as

proxies for those of real-world others and, therefore, less likely to undervalue the experience-sharing strategy. Indeed, the previous chapter identified that studies in Dodell-Feder and Tamir's (2018) meta-analysis of causal fiction effects had shown average larger effects of fiction-reading on experience-sharing than on mentalising or prosocial concern dimensions. Study 2 aimed to assess how far fiction-exposure relates to mentalising and experience-sharing strategies for empathic accuracy, and whether the likelihood of selecting one strategy over the other would vary as a function of fiction-exposure (based on Zhou et al.'s, 2017, approach). In so doing, it could help to establish whether the strategies deployed in understanding others in the real world, which have varying success rates, may be influenced by fiction, both in terms of the level of accuracy that they generate, and their perceived value. Specifically, the following hypotheses were tested:

H9: Fiction-exposure would positively predict empathic accuracy.

H10: There would be an interaction effect of fiction-exposure and empathic strategy on empathic accuracy, such that participants with high fiction-exposure would score higher on the empathic accuracy task, and this effect would be greater in the experience-sharing condition rather than the mentalising condition.

H11: Fiction-exposure would predict empathic strategy choice such that participants with higher fiction-exposure would show a higher likelihood of selecting experience-sharing rather than mentalising.

Study 2 aimed to replicate previous findings that fiction-exposure is positively associated with empathic accuracy by using a behavioural task measure that distinguished mentalising and experience-sharing inferencing strategies. It also aimed to explore the impact of contextual factors, the effects of valence and of different target experiencers. Considering the potential for an ingroup advantage (Matsumoto et al., 2009), the extent that participants identified with the target experiencers was also explored. Study 2 represents the first explicit

comparison of mentalising versus experience-sharing empathic strategies in the relationship between fiction-exposure and empathic accuracy.

4.1.3 Study Three: Effects of immersion in fiction on empathic accuracy.

Mixed findings from experimental studies of fiction effects on empathic accuracy have cast doubt on the hypothesis that reading a passage of fiction can enhance empathic skills. However, Dodell-Feder and Tamir's (2018) meta-analysis revealed a small but positive effect of fiction-reading on empathic accuracy task performance overall. The third and final study in the present research aimed to contribute to this literature via an experimental approach.

The mechanisms underpinning fiction effects remain unidentified (Dodell-Feder & Tamir, 2018). As argued in Chapter 3, immersion in stories, the experience of becoming "lost" (Nell, 1988, p. 8) or "transported" (Gerrig, 1993), is conceptually linked with empathy. Some studies have shown that the tendency to become immersed in narratives supports empathic abilities (e.g., Bal & Veltkamp, 2013; Johnson, 2012), although other studies have revealed no effect (e.g., Kidd & Castano, 2013). As outlined in the previous chapter, research in the adjacent field of narrative persuasion has demonstrated that manipulating immersion in a story impacts its persuasive effects, such that high immersion leads to more story-consistent beliefs after reading (Green & Brock, 2000). The present research aimed to identify whether immersion in a story would impact empathic accuracy and, if so, whether effects would hold across empathy dimensions.

4.1.3.1 Dimensions of empathy. As Dodell-Feder and Tamir (2018) pointed out, studies have tended not to examine which specific empathic skills may be improved by fiction-engagement and so the measures used in such studies have included a range of mentalising, experience-sharing and prosocial behaviour dimensions. The range of tasks have incorporated mental state content that is either cognitive (e.g., false belief tasks), affective (e.g., Emotion Attribution Task) or both (e.g., Yoni test), and they have generally required explicit responses

to questions. In Mumper and Gerrig's (2017) meta-analysis of correlational studies reviewed in the previous chapter, only one vignette task measured implicit empathic inferencing, and in Dodell-Feder and Tamir's (2018) meta-analysis of experiments also reviewed in Chapter 3, only the SST measured spontaneous mentalising (along with accuracy for explicit mental state content).

The third study in the current project aimed to examine how far reading a passage of fiction would contribute to both explicit and spontaneous inferences about cognitive and affective mental state content. To do so, two of the multifaceted tests used in previous studies were employed: the Yoni test and the SST (both described in Chapter 2). The Yoni test uses static vignettes and therefore requires mentalising based on visual cues. The SST, in contrast, incorporates questions about naturalistic story stimuli and may recruit both mentalising and experience-sharing inferencing strategies. Combined, these tests assess spontaneous empathic inferencing and explicit empathic accuracy for cognitive and affective mental state content. Both are capable of identifying variation among NT adults. As proposed in Chapter 2, using complex, naturalistic tasks and simpler tasks concomitantly enables a range of empathic accuracy dimensions to be measured without compromising construct validity. Specifically, the research aimed to test the following hypotheses:

H12: Participants who were immersed in a narrative during reading would show higher empathic accuracy for cognitive content after reading compared to participants who were not immersed.

H13: Participants who were immersed in a narrative during reading would show higher empathic accuracy for affective content after reading compared to participants who were not immersed.

H14: Participants who were immersed in a narrative during reading would show higher explicit empathic accuracy after reading compared to participants who were not immersed.

H15: Participants who were immersed in a narrative during reading would be more likely to make a spontaneous empathic inference after reading compared to participants who were not immersed.

4.1.3.2 Dimensions of immersion. In the previous chapter it was argued that immersion, like fiction and empathy, is multidimensional. Being immersed in a story may entail imaginative imagery, identification with characters and emotional engagement, and these dimensions may relate differentially to the mentalising, experience-sharing and concern components of empathy. It may be that fiction—or certain fictional content—induces one or more dimensions in particular and, in turn, these may influence different facets of empathy. Thus, Study 3 aimed to examine whether immersion in a story impacts empathic accuracy after reading, and to account for the potential contribution of its different dimensions.

4.2 General Methodology and Data Analysis Procedures

The present study aimed to contribute to the literature identifying relationships between lifetime exposure to fiction and empathic skills, as well as causal effects of narrative-engagement on empathic accuracy. The correlational and experimental approaches taken in the current research project are in line with previous studies of fiction effects on empathic domains described in Chapter 3, and the measures used were based on existing research. Established empathy scales and behavioural empathic accuracy tasks were selected (the IRI, self-report altruism scale, Yoni test and SST, all described in Chapter 2), and a version of Zhou et al.'s (2017) empathic accuracy task (see Chapter 2) was created due to unavailability of the original stimuli. In order to assess correlations with fiction-exposure, two versions of the ART were used (Acheson et al., 2008; Mar et al., 2006), and the latter was updated for the current study (Appendix D). There was no extant tool to assess lifetime exposure to different media, and a new version of the ART was created for this purpose (see Chapter 5 and Appendix A). Other questions (e.g., demographic questions, control and exploratory questions) were constructed using a range of response modes, including Likert-style scales,

text entry, ranking questions and slider scales. Materials for all studies were presented in Qualtrics unless otherwise stated (e.g., hard copy stimulus text in Study 3, Experiment 1).

Data were collected online (Study 1, Study 2 and Study 3 Experiments 2–4) and in a university lab (Study 2 stimuli; Study 3 Experiments 1 and 5). All data analyses were conducted using SPSS versions 23 and 24 (IBM Corp., Released 2015, 2016) and power analyses using SPSS syntax or G*Power (Faul et al., 2007). All studies received a favourable opinion from the Kingston University ethics committee and complied with the British Psychological Society's standards for the treatment of human participants.

4.3 Potential Implications

4.3.1 Implications for scholarship

The research presented in this dissertation took a multidimensional approach to investigating how fiction-engagement may benefit empathy. The findings could inform knowledge in three core areas: (i) scholarship on the nature of relationships between fiction and empathic abilities, (ii) approaches to the measurement of fiction-exposure and empathic accuracy, (iii) knowledge of the processes involved in fiction effects.

First, the primary aim of this research was to contribute to knowledge of relationships between fiction and empathic abilities as, to date, findings have been mixed. Previous studies have incorporated a range of approaches, which may account for some heterogeneity in the extant findings. The multidimensional approach taken here aimed to provide a more granular perspective on these relationships, and to foreground approaches to measurement. Across the whole project, a variety of stimuli enabling access to a range of empathy components— including cognitive and affective mental state content, explicit and spontaneous mentalising and experience-sharing processes—were used in order to scope and clarify the potential effects of fiction-engagement. To the author's knowledge, the present research represents the first to examine correlations between exposure to different fiction media and empathic

abilities, to examine relationships between fiction-exposure and mentalising versus experience-sharing strategies using a behavioural approach, and to investigate the immediate effects of manipulating immersion in a story on a range of cognitive and affective behavioural empathic ability dimensions.

Second, two measures were developed and produced as part of this research: an adaptation and extension of the ART which incorporated three fiction media formats, and a version of Zhou et al.'s (2017) empathic accuracy tool, which facilitated the comparison of mentalising versus experience-sharing processes. These paradigms may be further developed in future research. For example, the empathic accuracy task design, which diverged from Zhou et al.'s method in that the same set of target experiencers was examined across all participants (whereas Zhou et al. varied experiencers between participants), could be developed to facilitate measurement of the moderating effects of demographic differences between target experiencers (and their relationships to participant demographics). The new ART measure could be extended to include other fiction media scales (e.g., video game-play) in order to support the development of fiction effects models that reflect the multiplicity of the fiction landscape.

Third, clarifying the role of immersion would develop knowledge of the conditions in which fiction effects occur. This knowledge could contribute to new approaches to classifying fiction for the purposes of experimental research; for example, by indexing attributes that contribute to immersion. In turn, this could help to develop theoretical models of the processes through which fiction affects empathy dimensions. It could also impact empathy research more broadly: in addition to forming the stimuli in studies of fiction-engagement, fictional narratives are also regularly employed in test protocol aimed at assessing empathic accuracy (e.g., Dodell-Feder et al., 2013; Dziobek et al., 2006; Happé, 1994). They have been useful in studies with NT adults because they are less susceptible to ceiling effects than

traditional tasks (for an example of ceiling effects using a theory of mind test in a study from the field of fiction effects, see Kidd & Castano, 2013). However, individual differences in immersion tend not to be accounted for in such studies. If immersion in fiction impacts empathic accuracy, variance in scores obtained via fiction-based tasks may be confounded by differences in immersion. If this is the case, empathy researchers employing stories as stimuli should consider partialling out variance in empathy scores caused by individual differences in immersion or the trait tendency to become immersed (e.g., IRI fantasy subscale). Beyond a contribution to the field of fiction effects, then, this research may also have implications for the measurement of empathy in other fields.

4.3.2 Social implications

Empathic abilities are essential for making sense of the social world, maintaining interpersonal relationships and generating community cohesion. The real-world implications of the research are, therefore, potentially far-reaching. As discussed in Chapter 1, anecdotal evidence for the positive social impact of fiction already exists through reports from book clubs, theatre in education initiatives, and role-play in corporate, medical, social care and criminal justice contexts, which have indicated that fiction-based interventions can enhance social functioning. Studies showing positive effects of fiction on social outcomes have provided some empirical support for such claims, but because results have been mixed, and key studies have failed to replicate, the field has yet to reach concrete conclusions about whether, and under what conditions, fiction impacts empathy.

If the current project found effects of fiction on empathic accuracy with NT adults, it would support the idea that empathic skills can be trained, even at the upper echelons, through fiction-based interventions. By investigating the conditions in which such effects may occur, including different fictional themes, formats and levels of immersion, and through examining different dimensions of empathy, the current project could provide insight into how best to

exploit fiction as a tool for developing empathic and prosocial tendencies. This information may also lay foundations for future interventions with groups that show characteristic deficits in specific empathic dimensions (e.g., mentalising in ASD and experience-sharing in psychopathy; Freedman et al., 2013). Consequently, this research could help to bridge the gap between anecdotal and scientific accounts of the value of literary arts-based social initiatives.

4.4 Project Hypotheses

Table 2 presents the full set of hypotheses addressed across three studies.

Table 2

Research Hypotheses and Exploratory Aims of Studies 1–3.

	Hypotheses
Study 1	
H_1	Fiction-exposure via print, film and plays combined would positively predict perspective-taking.
H_2	Fiction-exposure via print, film and plays combined would positively predict fantasy.
H_3	Fiction-exposure via print, film and plays combined would positively predict empathic concern.
H_4	Fiction-exposure via print, film and plays combined would positively predict altruism.
	Further exploration: The research aimed to explore relationships between (i) exposure to fiction via print, film and plays and empathic abilities (perspective-taking, empathic concern, fantasy and altruism), (ii) preferences for media formats and thematic genres, and empathic abilities, (iii) experience of acting and empathic abilities.
Study 2	
H_5	There would be an effect of empathic strategy on empathic accuracy, with participants that used the experience-sharing strategy showing higher empathic accuracy compared to participants that used the mentalising strategy.
H_6	There would be an effect of valence on empathic accuracy.
H_7	There would be an effect of experiencer on empathic accuracy.
H_8	There would be an interaction effect of strategy, valence and experiencer on empathic accuracy.
H_9	Fiction-exposure would positively predict empathic accuracy.
H_{10}	There would be an interaction effect of fiction-exposure and empathic strategy on empathic accuracy, such that participants with high fiction-exposure would score higher on the

	empathic accuracy task, and this effect would be greater in the experience-sharing condition rather than the mentalising condition.
H_{11}	Fiction-exposure would predict empathic strategy choice such that participants with higher fiction-exposure would show a higher likelihood of selecting experience-sharing rather than mentalising.
	Further exploration: As facial expressions recognition can be affected by an ingroup advantage (Matsumoto et al., 2009), the research aimed to examine how far identification with experiencers would impact accuracy scores.

Study 3

H_{12}	Participants who were immersed in a narrative during reading would show higher empathic accuracy for cognitive content after reading compared to participants who were not immersed.
H_{13}	Participants who were immersed in a narrative during reading would show higher empathic accuracy for affective content after reading compared to participants who were not immersed.
H_{14}	Participants who were immersed in a narrative during reading would show higher explicit empathic accuracy after reading compared to participants who were not immersed.
H_{15}	Participants who were immersed in a narrative during reading would be more likely to make a spontaneous empathic inference after reading compared to participants who were not immersed.
	Further exploration: In order to assess the extent that different facets of immersion are impacted by fiction, and associated with empathy, three dimensions of immersion would be measured using an imaginative transportation scale, an emotion scale and an identification with character scale.

4.5 Summary and Moving Forward

Taking a multidimensional approach to the study of relationships between fiction and empathic abilities could develop knowledge of the potential benefits of fiction-engagement, and influence scholarship in the field of fiction effects as well as in empathy research. This chapter has provided an outline of three studies which aimed to contribute to these areas of knowledge and established specific hypotheses for each. The next three chapters (Chapters 5, 6 and 7) present the research conducted. Each chapter includes a brief abstract, introduction, method, results and discussion, and although each contains some signposting to ideas

presented in other chapters, they can be read as standalone research reports. Chapter 8

synthesises the full set of findings through a general discussion, identifies the strengths and

limitations of the project, and outlines avenues for future research.

Chapter 5: Relationships between Fiction Media, Genre, and Empathic Abilities

Fiction enables readers to simulate the social experiences of characters. Research using ARTs has indicated that fiction print-exposure positively relates to aspects of empathy and may promote prosocial behaviours (Mumper & Gerrig, 2017). Whether associations hold across different media formats and thematic genres remains unclear. This study took a multidimensional approach to both fiction-exposure and empathic abilities. Specifically, it aimed to replicate previous findings that lifetime fiction-exposure positively predicts empathic abilities. It aimed to extend this literature by examining relationships between exposure to fiction through reading and non-reading formats, mentalising and experience-sharing dimensions of empathy, and the tendency to behave prosocially. Further, it aimed to explore relationships between people's preferences for different fiction-engagement channels, genres, and empathic abilities. Participants ($N = 404$) completed a multidimensional task measure of fiction media-exposure designed for the study and answered questions about their fiction-engagement, empathic and altruistic tendencies. Results showed divergent associations between fiction format, genre, and empathic abilities, and fiction media-exposure positively predicted the tendencies to become imaginatively absorbed in narratives and to help others. Exposure to fiction via print positively predicted self-report altruism and fantasy. Fiction formats and thematic genres were differentially related to adults' empathic abilities. The current chapter reports this correlational study, beginning with a recap of the relevant literature.

5.1 Introduction

The ability to understand others and respond appropriately to their needs is central to the formation of successful interpersonal relationships and cohesive communities (Castano, 2012; Paal & Bereczkei, 2007). Recall that prosocial, helpful behaviours involve both empathic and altruistic personality dimensions (Penner et al., 1995) and that empathy can lead to altruism

(the willingness to help others with no expectation of reward, Penner et al., 1995) by enabling

perceivers to interpret others' mental states (Batson et al., 1981). As discussed in Chapter 3,

these skills typically develop in childhood (e.g., Perner & Wimmer, 1985) and so studies have

tended to focus on children or groups with characteristic deficits. However, empathic abilities

vary between NT adults and can change through the lifespan (Duval et al., 2010; Happé et al.,

1998; Maylor et al., 2002). Research has begun to address how adults' empathic skills and

prosocial tendencies may be enhanced.

5.1.1 Fiction effects on empathic abilities

Fiction-reading may cultivate social understanding and increase prosocial behaviour via

the transmission of social knowledge, or because fiction comprehension involves the same

cognitive processes as real-world events (Mar et al., 2009; see also Gerrig, 1993). The

"process" view (Mar et al., 2009) that readers foster their empathic abilities by mentally

simulating the social experiences depicted in fiction (Mar & Oatley, 2008; Oatley, 1999,

2016), has received support from three strands of research: first, neuroscience has revealed

that readers' brains show activity in areas associated with the experiences they read about

(e.g., Speer et al., 2009; Wallentin et al., 2011) and that narrative fiction activates the brain's

empathy network (Hsu et al., 2014). Second, cognitive experiments have demonstrated

positive effects of fiction-reading on both mentalising (explicit reasoning about thoughts and

emotions) and experience-sharing capacities (vicariously experiencing another's inner state,

e.g., Coll et al., 2017; and see Dodell-Feder & Tamir's, 2018, meta-analysis), as well as

prosocial behaviour (e.g., Johnson, 2012; Koopman, 2015), and attitude change (e.g., Appel &

Richter, 2007; Green & Brock, 2000). Although attitude change does not necessarily

implicate empathic or prosocial processes, empathising with an individual's story can

improve attitudes towards their social group (Batson et al., 1997). If the same mechanisms are

involved in real-word and fiction-induced empathy, empathy for fictional characters—which

serve as proxies for real-world others—could lead to real-world altruism (see Batson et al., 1981; Klimecki, Mayer, Jusyte, Scheeff & Schönenberg, 2016; though see Keen, 2007). Third, correlational studies have shown that while exposure to nonfiction print is positively associated with empathic understanding, correlations tend to be larger for fiction (Mumper & Gerrig, 2017). While expository nonfiction can feature social themes and humanlike agents, fiction is distinguished by its complex characterisation and narrative structure, which may recruit, and consequently strengthen, the psychological mechanisms involved in intersubjectivity (Mar & Oatley, 2008).

The general hypothesis that fiction immediately enhances people's capacity to understand real-world others has been called into question. Studies have failed to reproduce Kidd and Castano's (2013) finding that reading literary fiction immediately enhances mentalising, finding no effect compared to nonfiction, popular fiction and no reading conditions: replication efforts by Camerer et al. (2018), Panero et al. (2016), Panero et al. (2017), and Samur et al. (2018) failed, whereas those by Kidd and Castano, (2018b), and van Kujik et al. (2018) yielded mixed results. These findings have raised doubts about causation, although Dodell-Feder and Tamir's (2018) meta-analysis indicated a small effect of fiction on empathic abilities (Hedge's g =.15–.16). However, Kidd and Castano's (2013) finding that fiction-exposure (measured using an ART) was a moderate predictor of mentalising ability (Experiments 1, 4 and 5, ω_{p2} = .13, .15, and .07 respectively; see Cohen's, 1988, rules of thumb for effect size) was replicated in both studies. Therefore, causal effects seem unstable (cf. Kidd & Castano's, 2018a, response to the failed replication attempts), though evidence for the relationship between lifetime exposure to fiction and empathic abilities appears robust.

5.1.2 Modes of fiction-engagement

Studies of associations between fiction and empathic abilities have primarily focused on reading. However, fiction is not only read in print; it can be experienced onscreen and at the

theatre, as well as through first-person video games and live acting or roleplay. Evidence suggests that engagement modality may impact empathic abilities: viewing filmed narratives has been shown to enhance mentalising in adults (Black & Barnes, 2015a), and to predict the ability in children (Mar et al., 2010), and live theatre has been found to alter audiences' social attitudes (Heide et al., 2012; Stephens-Hernandez et al., 2007; Valente & Bharath, 1999). Interactive video game play has been linked to increases in prosocial behaviour (Gentile et al., 2009), and acting lessons have been found to improve social behaviours in children (Schellenberg, 2004) and may increase empathic understanding in children and adolescents (Goldstein & Winner, 2012). Adolescent acting students (Goldstein et al., 2009) and professional adult actors (Nettle, 2006) tend to score higher than matched controls on self-report and task-based empathic ability measures. These findings suggest that processes involved in engaging with fictional content, rather than those specific to reading, are linked to empathic skills. The present study aimed to examine whether there is a cumulative effect of different forms of fiction-engagement on empathic abilities, and how far effects vary across modes of engagement.

5.1.3 Literary and genre-fiction

Despite evidence that fiction-exposure accounts for variance in empathic abilities above and beyond nonfiction-exposure (Mumper & Gerrig, 2017), there has been relatively little research examining the thematic and stylistic features of fiction that evoke responses to stories (e.g., Valkenburg, Peter & Walther, 2016). Kidd and Castano (2013) proposed that the relationship between fictional narratives and the ability to understand others is unique to literary fiction (their study used acclaimed and canonical texts). They argued that whereas popular fiction is generally formulaic and predictable, literary prose entails active, "writerly" comprehension, requiring readers to interpret complex narratives by establishing characters' motivations, thoughts and emotions, thus engaging their capacities for mentalising. Across a

series of five experiments (reviewed in Chapter 3), they assigned participants to read segments of either literary, popular or nonfiction prose, and then assessed their mentalising abilities using the RMET. They found that participants in the literary fiction condition tended to score higher on the task. Subsequent research has supported this finding. Using a pre- and post-test design, Pino and Mazza (2016) showed improved mentalising after reading literary compared to science-fiction and nonfiction texts, and reading a text high in literary foregrounding (striking textual features) compared to non-literary versions of the same text has been found to lead to higher empathic understanding for emotional experiences (e.g., of grief; Koopman, 2016). Using onscreen stimuli, Black and Barnes (2015a) found that participants who viewed award-winning TV dramas showed higher mentalising abilities compared to those who viewed documentaries, and recent studies have indicated positive relationships between familiarity with literary fiction and mentalising ability (De Mulder et al., 2017, Kidd & Castano, 2017a).

Recent research, however, has raised doubt about the unique value of literary fiction in providing "grist for the mills" (Zunshine, 2006, p. 16) of empathic abilities. Exposure to the romance genre, for example, has been shown to positively predict mentalising (Fong et al., 2013). It seems unlikely that romance contains more literary features than other genres or that romance exposure is particularly associated with acclaimed literature; on the contrary, it has been found that romance readers are less likely to appreciate figurative language, multiple plotlines and perspectives than literary fiction-readers, Miesen, 2004). Rather, the genre's emphasis on social interactions may potentiate empathic development. Furthermore, popular genre-fiction is associated with moral reasoning (Black et al., 2018), reading about certain social themes (e.g., depression) can motivate prosocial behaviour (Koopman, 2015), and fiction-generated suspense has been linked to brain regions involved in social cognition

(Lehne et al., 2015). If fiction effects are associated with specific thematic content, relationships between thematic genres and empathic abilities are likely to vary.

5.1.4 Towards the multidimensional approach

Fiction is multidimensional and can be classified in terms of thematic genre as well as media format. Empathic abilities, too, are multifaceted: understanding and responding prosocially towards others involves mentalising and experience-sharing components (Zaki & Ochsner, 2012) and extant research has documented a lack of correlation among measures (Davis, 1980; Ickes, 1997; Mar et al., 2006). Studies assessing relationships between fiction and empathic ability measures have yielded mixed findings; for example, Kidd and Castano (2013) found that reading a passage of literary fiction led to higher scores on the RMET, the most common measure used in the field (e.g., Black & Barnes, 2015a, 2015b; Fong et al., 2013; Kidd & Castano, 2013, 2017b; Mar et al., 2006; or see the meta-analyses by Dodell-Feder & Tamir, 2018; Mumper & Gerrig, 2017). Pino and Mazza (2016) did not replicate this result but showed improvements on two other tests of mentalising (a false-belief task and a facial emotion recognition test) and no improvement on experience-sharing measures. Inconsistencies across the literature may reflect different empathic processes probed by a range of measures, as well as the employment of broad, heterogeneous fiction classifications.

As described in Chapter 3, Author Recognition Tests (ARTs; Stanovich & West, 1989) measure familiarity with fiction authors as a proxy for lifetime exposure to fiction-in-print, and they have been widely used to assess relationships between fiction-exposure and empathic abilities. There currently exists no equivalent tool to assess exposure to fiction via different media presentations, which would help to reflect the range of formats that showcase fictional storytelling. Despite actors' portrayals of characters being widely available onscreen and at the theatre, the extent to which the processes involved in interpreting these resemble those involved in interpreting character depictions in printed novels has received little

empirical attention (Goldstein & Filipe, 2018). Thus, it cannot be taken for granted that fiction effects on empathic abilities would be the same across different media formats. The aims of this study were to (i) replicate previous findings that fiction-exposure positively relates to empathic abilities, (ii) extend the literature by assessing the cumulative effect of exposure to different fiction formats through the construction of a multidimensional measure (testing Hypotheses 1- 4 articulated in Chapter 4), and (iii) explore how far relationships between modes of engagement with fiction, thematic genres and empathic abilities would vary (the exploratory analysis articulated in Chapter 4, section 4.1.1).

5.2 Method

5.2.1 Participants

Participants were recruited via a university research participation scheme (24%), a research participant recruitment website (Call for Participants, 10%), or followed links to the Qualtrics survey posted to university and researcher accounts on social media sites (Facebook and Twitter, 66%). Consequently, respondents were either undergraduate Psychology students or members of the public interested in research participation. The survey was closed at 405 completed responses. One participant was excluded for having fewer than three years' English fluency (one year), resulting in total $N = 404$ (81% females, $M_{age} = 36.5$, $SD_{age} = 13.77$, age range 17-74). Participants were native English speakers (85%) or reported a minimum of three years of fluency ($M = 13.4$, $SD = 8.6$). Average time spent in post-compulsory education was 6.2 years ($SD = 3.68$). Participants were compensated with university course credit, if relevant, and were given the opportunity to enter a prize draw to win a £100 online shopping voucher. The study was given a favorable opinion by the Faculty Research Ethics Committee at Kingston University London and complied with the British Psychological Society's standards for the treatment of human participants.

5.2.2 Materials

5.2.2.1 Fiction Media-exposure Test. The Fiction Media-exposure Test (FMET) is a measure of exposure to fiction adapted for this study from previous ARTs (Acheson et al., 2008; Mar et al., 2006). The original ART (Stanovich & West, 1989) provides a well-validated, proxy measure of fiction-exposure through the recognition of author names. It is assumed that participants recognise author names due to having read their books or through browsing related authors in libraries, bookshops and online. Because familiarity with fiction may be considered socially desirable, the test contains both real fiction authors (e.g., James Joyce) and plausible foils (e.g., Robert Tierney) to deter guessing.

The FMET was designed to provide a measure of fiction-exposure along three independent dimensions: print, film and play-exposure (Appendix A). The play-exposure dimension was used to examine exposure to performed plays (theatre), rather than to those that may be read, for example, on literature courses. Therefore, names were selected from two lists of significant 20th century plays generated by the National Theatre (n.d.) and the New York Theater (2003). In line with earlier ARTs, name recognition was treated as a proxy measure of exposure to plays through theatre attendance, familiarity with playwrights, or via advertisements in venues and online. Similarly, films were Golden Globe Best Motion Picture award winners and nominees (Hollywood Foreign Press Association, n.d.) from 1953 to 2016. Recognition is assumed to reflect knowledge of films through viewing at cinemas, on DVD or via streaming apps, familiarity with related works, or advertising. The print-exposure scale contains 30 fiction author names featured in an earlier ART (Acheson et al., 2008). Names were selected to represent authors of printed fictional prose that showed a range of recognition rates in Acheson et al.'s (2008) study: recognition rates for the 30 names selected for the FMET varied from 5%-99% of participants compared to Acheson et al.'s (2008) recognition rates of 2%-99%. The FMET comprised 30 names in each dimension, alongside

25 foils in the print (author recognition) scale, and 15 in each of the film and play dimensions (foil-selection tends to be low on ARTs; Fong et al., 2013; Mar et al., 2006).

To correct for indiscriminate responding, FMET scores were computed as the number of correct names selected minus the number incorrect foils selected within each scale. Internal consistency was very good with the present sample, indicated by a Cronbach's alpha value of .97 for the full measure and .92, .91 and .93 for the print, film and play scales respectively. Item deletion did not improve internal consistency except in the case of one play name (*Six Degrees of Separation*); as the increase was not substantial ($\alpha = .94$), the item was retained in the analysis.

5.2.2.2 Fiction preferences. Fiction-exposure does not reflect depth or quality of engagement, and some media may be less accessible than others (e.g., due to location and the availability of resources). Therefore, participants were asked to rank-order their favourite to least favourite media from four options regardless of external factors such as time, cost and convenience. Three options, "read a novel", "watch a film", and "watch a play", paralleled the FMET dimensions. Watching television represents the most popular leisure activity in the US (Bureau of Labor Statistics, 2017) and UK (Seddon, 2011), and so a fourth option, "watch a TV show/series", was included. Responses to this question triggered a bespoke element, with participants subsequently asked to rank-order the reasons they tended to select their preferred three media formats. Seven options reflected information and education ("interest in creatives", "learning"), enhancing social interactions ("to discuss it with others", "it's something I should have seen/read"), escapism ("to forget about things") and affect ("mood improvement", "mood indulgence"; see McQuail, 2010). The results for these follow-up questions are reported in Appendix B.

Enjoyment of six genres was measured to test associations between thematic content and empathic abilities. Drama, comedy, crime/thriller, romance, and experimental/postmodern

were selected as diverse examples of fiction genres that could apply to each media format. Additionally, the factual/documentary genre was included based on previous findings that, compared to fiction, reading or viewing nonfiction is associated with lower scores on mentalising tests (Black & Barnes, 2015a, 2015b; Kidd & Castano 2013; Mar et al., 2006). Participants indicated enjoyment of each genre on 5-point Likert scales from "not at all" to "a great deal".

5.2.2.3 Acting experience. Experience of participation in fiction through acting was assessed using a single-item 5-point scale with answers ranging from "no experience" to "highly experienced (current professional)".

5.2.2.4 Empathic abilities. In line with Penner et al. (1995), who provided a two-factor solution for measures of prosociality using a battery composed of altruism and empathy scales, altruism was measured using a truncated version of Rushton et al.'s (1981) self-report altruism scale, and empathy was measured using three scales from the Interpersonal Reactivity Index (IRI; Davis, 1980).

The IRI asks respondents to indicate agreement with statements on 5-point Likert scales across four independent 7-item dimensions: perspective-taking, empathic concern, fantasy and personal distress. Recall that Mumper and Gerrig's (2017) meta-analysis revealed significant correlations between fiction-reading and all IRI dimensions except for the personal distress subscale. This scale measures self-oriented responses to others' needs—i.e., the desire to reduce one's own distress—which can lead to avoidance of empathic behaviours, distinguishing it from other IRI dimensions (Baron-Cohen & Wheelwright, 2004; Batson & Shaw, 1991; Davis et al., 1999; Penner et al., 1995; Singer & Lamm, 2009) and so it was not used in this study. The perspective-taking scale measures the ability to mentalise about other people's points of view (e.g., "I try to look at everybody's side of a disagreement before I make a decision"). The fantasy scale assesses the tendency to become absorbed in imagined

worlds and share in the experiences of predominantly fictional characters (e.g., "after seeing a play or movie, I have felt as though I were one of the characters") and represents the experience-sharing dimension. The empathic concern scale measures prosocial concern towards others (e.g., "I often have tender, concerned feelings for people less fortunate than me"). Cronbach's alpha (α) values for the IRI with this sample were .77 for empathic concern, .81 for perspective-taking and .81 for fantasy.

The self-report altruism scale requires participants to indicate the frequency with which they have carried out acts that refer to helping acquaintances, strangers, and charities (e.g., "I have given money to a stranger who needed it or asked me for it"), on a 5-point scale from "never" to "very often". Of the original 20 items, nine covering a range of helpful acts toward acquaintances, strangers and charity were selected and modified for use in the present study. The scale was truncated to reduce the number of questions participants were required to answer in the study, and internal consistency using this approach was satisfactory ($\alpha = .70$).

5.2.2.5 Control variables and demand characteristics. Self-report data were gathered on age, gender, English fluency and years of education as these may impact exposure to fiction media or empathic abilities (for instance, gender differences have been shown in empathic abilities; Davis, 1980). Irrelevant items were integrated with the IRI and self-report altruism scale to mask the purpose of the research (e.g., "I get little enjoyment from cleaning my home"). Participants were informed that points would be deducted for foil-checking on the FMET, and one point was deducted for each foil selected. Respondents could skip or select "I don't know" to any of the questions.

5.2.3 Procedure

The study was administered using Qualtrics and took approximately 15-minutes to complete. After providing consent, participants answered the demographic, and fiction media and genre preferences questions. The primary measures, the FMET and empathic ability

questions (altruism and IRI scales) were then administered in a random order. Apart from the demographic questions, all measures were internally randomised. Participants were fully debriefed online.

5.2.4 Data analysis

Reliability analyses were conducted on the FMET, IRI and altruism scales and Cronbach's alpha values are reported above. The data showed heteroscedasticity and non-normality across the FMET, IRI and altruism scales (assessed with visual inspection of plots and the Shapiro-Wilk test), and the presence of outliers on all scales except fantasy and print-exposure. Therefore, non-parametric and robust statistical tests were employed: the Mann Whitney-U Test was used to examine effects of gender and language, weighted least squares multiple regression analyses were used to predict empathic abilities from fiction-exposure and control variables, and Spearman's rho (r_s) was used for raw correlations. Partial correlations were conducted using Pearson's *r* following rank transformations on the data, and controlled for age, gender and education (Appendix B). All correlations were computed using pairwise deletion, and confidence intervals were bias-corrected and accelerated using Preacher and Hayes' (2008) bootstrapping procedure ($N = 5000$).

5.3 Results

5.3.1 Demographic variables

Age and education were significantly associated with all fiction-exposure dimensions and altruism, all ps < .001. There were significant effects of language on print, film and play-exposure, all ps < .001, while effects on empathic abilities were non-significant. Females scored higher than males on all IRI scales, though there was no effect of gender on altruism. Males recognised more film names than females, but males and females did not differ on the print or play-exposure dimensions of the FMET (gender differences are reported in Table 3).

Table 3

Gender Comparisons for IRI, Self-report Altruism Scale and FMET.

	Females		Males			
	Mdn	*SE*	*Mdn*	*SE*	*U*	*p*
Perspective-taking	26	.50	23	.69	14878.50	.001
Empathic concern	29	.32	26	.73	16241.50	< .001
Fantasy	25	.41	23	1.13	14419.50	.004
Altruism	25	.49	25	.79	12446.50	.498
Print-exposure	12	.60	12	1.08	11003.00	.345
Film-exposure	18	.51	20	.85	9082.00	.002
Play-exposure	6	.58	8.5	1.25	10109.50	.051

Note. N = 401 (329 participants identified as female and 72 as male).

5.3.2 Fiction media-exposure

Means, standard deviations and inter-scale correlations for the FMET, IRI and altruism scales are presented in Table 4. Participants recognised more film names ($M = 17.9$, $SD = 6.79$), than print authors ($M = 11.92$, $SD = 6.95$), or plays ($M = 7.47$, $SD = 6.69$), $F(2, 806) = 922.20$, $p < .001$, $\eta_2 = .70$, $p < .001$, for all comparisons. Foil-checking on the FMET was low: 95% of participants selected fewer than five foils across all three scales. There were strong positive associations among the print, film and play recognition scales, as well as among the IRI and altruism dimensions, all ps < .001. Additionally, fantasy positively correlated with print-exposure, and altruism with all fiction-exposure scales, all ps < .01. Partial correlations ($df = 393$) revealed a similar pattern of results (Appendix B): associations reached significance between fantasy with play, $r = .15$, $p = .003$, 95% CI [.05, .24], and film-exposure, $r = .15$, $p = .002$, 95% CI [.05, .26], though the correlation between altruism and film-exposure did not reach significance.

Table 4

Means, Standard Deviations and Raw Inter-Scale Correlations for FMET, IRI and Self-Report Altruism Scale.

	M	*SD*	1	2	3	4	5	6	7
1. Print-exposure (α = .92)	11.92	6.95	-	.74 *** [.69, .79]	.80 *** [.75, .83]	.29 *** [.18, .36]	.05 [-.05, .14]	-.01 [-.11, .08]	.13 ** [.04, .22]
2. Film-exposure (α = .91)	17.90	6.79		-	.79 *** [.75, .83]	.24 *** [.05, .21]	.05 [-.05, .14]	.02 [-.08, .15]	.08 [-.01, .19]
3. Play-exposure (α = .93)	7.47	6.69			-	.29 *** [.21, .38]	.04 [-.05, 14]	.001 [-.06, .07]	.09 [-.01, .19]
4. Altruism (α = .70)	25.12	5.60				-	.30 *** [.21, .39]	.27 *** [.17, .36]	.13 ** [.03, .22]
5. Perspective-taking (α = .81)	24.46	5.27					-	.55 *** [.48, .62]	.30 *** [.21, .39]
6. Empathic concern (α = .77)	27.75	4.65						-	.38 *** [.29, .47]
7. Fantasy (α = .81)	24.30	5.88							-

Note. N = 404, * p < .05, ** p < .01, *** p < .001, 95% bias-corrected and accelerated confidence intervals are presented in brackets.

5.3.2.1 Regression analyses. Regression models were conducted to predict each of the empathic abilities using FMET scores (scale), gender (categorical), age (scale), education (ordinal converted to scale for multiple regression) and language (categorical).

5.3.2.1.1 Assumption tests. There was independence of observations for all models, indicated by Durbin Watson statistics of near 2 (values were 1.93, 2.09, 1.94 and 1.93 for altruism, fantasy, empathic concern and perspective-taking respectively). Correlations among predictors were < .7 for all models, indicating no problem with multicollinearity. Further, there were satisfactory Tolerance (> .1) and Variance Inflation Factor (VIF) values (< 10). All models met the assumption of linearity, assessed via visual inspection of studentised residuals against predicted values. Inspection of partial regression plots showed linearity between each IV and DV, and the data were normally distributed. In the model for fantasy, there was a slight negative skew, but this was not substantial and did not require transformation of the data. There were no outliers in the model for fantasy. There were three outliers in the altruism data (inspection of studentised residuals identified values that were -3.2, +3.47 and +3.19 *SD*s from the mean), three in the empathic concern data (-3.25 -3.46 and -3.85 *SD*s), and one in the perspective-taking data (-3.25 *SD*s), but there were no high leverage or highly influential points in any case (all Cook's Distance values < 1; Cook & Weisberg, 1982) and so all of these values were retained. Visual inspection of a scatterplot of the studentised residuals against the predicted values indicated some heteroscedasticity, hence multiple least squares regression models were used. These revealed the same pattern of results as standard multiple regression models.

5.3.2.1.2 Weighted least squares regression models. Fiction-exposure (the combined FMET scales) and control variables (gender, age, education and language) significantly predicted 4% variation in fantasy, 6% of empathic concern, 4% of perspective-taking and 10% of altruism, all ps < .01. Coefficients (*B*) and model fit statistics are reported in Table 5.

Gender was the only significant contributor to the models for empathic concern and perspective-taking, both $ps < .001$, whereas fiction-exposure, age and gender were significant contributors to fantasy, all $ps < .05$, and fiction-exposure and age to altruism, both $ps < .01$. Due to correlations between the IRI and altruism measures, regression analyses were also conducted with the inclusion of each of the three remaining empathic ability scales as predictors, which revealed the same pattern of results.

Table 5

Predictors of Empathic Abilities.

	Altruism	Fantasy	Empathic concern	Perspective taking
Coefficients (*B*)				
Constant	18.88 ***	17.43 ***	21.32 ***	19.14 ***
	[15.07, 22.69]	[13.23, 21.41]	[18.32, 24.33]	[15.82, 22.47]
Fiction-exposure	0.05 **	0.06 **	0.02	0.03
	[.01, .08]	[.02, .10]	[-.01, .05]	[-.001, .07]
Age	0.08 ***	-0.05 *	0.01	-0.03
	[.03, .13]	[-.10, -.001]	[-.03, .05]	[-.08, .01]
Gender	0.70	2.55 **	2.85 ***	2.38 ***
	[-.79, 2.20]	[.94, 4.16]	[1.74, 3.95]	[1.17, 3.59]
Education	0.01	0.08	-0.003	0.09
	[-.14, .17]	[-.08, .25]	[-.11, .10]	[-.04, .21]
Language	0.15	1.27	0.29	0.46
	[-1.44, 1.74]	[-.38, 2.90]	[1.08, 1.65]	[-1.05, 1.97]
Model fit				
R_2	.10	.05	.07	.05
adj. R_2	.10	.04	.06	.04
F	9.35	4.32	5.70	3.88
p	< .001	.001	< .001	.002

Note. $N = 399$; *$p < .05$, **$p<.01$, ***$p < .001$; 95% bias-corrected and accelerated confidence intervals are presented in brackets.

5.3.3 Genre and media preferences

Comedy was the most popular genre (38% of participants reported the highest enjoyment rating), followed by crime/thriller (37%), factual/documentary (31%), drama (30%), romance (13%) and experimental/postmodern (5%). Raw correlations for genre enjoyment, media preference and empathic abilities are presented in Table 6. Significant associations between genre preferences and empathic abilities were all positive: fantasy was associated with all genres except crime/thriller and factual/documentary, all ps < .01, perspective-taking and empathic concern with all genres except crime/thriller, and altruism with experimental/postmodern, factual/documentary and comedy, all ps < .05. Partial correlations revealed a similar pattern of results (Appendix B).

Watching a film was the preferred media option (57% of participants ranked it as their top or second preference), followed by reading a novel (53%), watching TV (52%), and watching a play (38%). Preference for reading novels positively correlated with fantasy and altruism, both ps < .001, and preference for plays with altruism, p < .001, and empathic concern, p < .05. In contrast, preference for film negatively correlated with altruism, p < .01, and preference for TV with altruism, p < .001, perspective-taking, p < .01, and empathic concern, p < .05. Partial correlation results were similar (Appendix B); however, the positive relationship between altruism and reading novels, and the negative relationship between fantasy and watching TV were marginal, p = .093 and p = .055, respectively.

Results revealed both positive and negative relationships among media and genre preferences (df = 351): there were positive associations between preference for plays and enjoyment of drama, r_s = .16, p = .003, 95% CI [.06, .25] and experimental/postmodern, r_s = .24, p < .001, 95% CI [.14, .34]; and negative associations between preference for novels

Table 6

Correlations between Genre Enjoyment, Media Preferences, and Empathic Abilities.

	Altruism	Perspective-taking	Empathic concern	Fantasy
Genre Enjoyment				
Drama	.08 [-.02, .18]	.11 * [.004, .20]	.18 *** [.08, .29]	.27 *** [.17, .36]
Romance	-.06 [-.17, .04]	.15 ** [.05, .24]	.27 *** [.16, .35]	.21 *** [.11, .31]
Crime/Thriller	.04 [-.07, .14]	.07 [-.04, .17]	.06 [-.05, .16]	.06 [-.05, .16]
Experimental/Postmodern	.14 ** [.03, .24]	.18 *** [.08, .28]	.17 ** [.06, .27]	.14 ** [.04, .24]
Comedy	.13 * [.03, .23]	.13 * [.03, .23]	.19 *** [.08, .29]	.18 *** [.08, .28]
Factual/Documentary	.13 * [.02, .23]	.13 * [.03, .23}	.13 * [.03, .23]	.08 [-.03, .19]
Media Preferences†				
Novels	.18 *** [.07, .27]	.08 [-.02, .18]	.04 [-.06, .14]	.16 *** [.07, .25]
Film	-.13 ** [-.23, -.03]	-.04 [-.13, .06]	-.04 [.05, -.13]	-.08 [-.18, .02]
Plays	.21 *** [.12, .31]	.08 [-.02, .18]	.11 * [.02, .21]	-.01 [-.11, .08]
TV	-.26 *** [-.35, -.17]	-.13 ** [-.23, -.03]	-.12 * [-.22, -.03]	-.09 [-.18, .01]

Note. $N = 353$, †$N = 403$, * $p < .05,$ ** $p < .01,$ *** $p < .001$, 95% bias-corrected and accelerated confidence intervals are presented in brackets.

and comedy, $r_s = -.10$, $p = .05$, 95% CI [-.21, -.002], and preference for TV and experimental/postmodern, $r_s = -.22$, $p < .001$, 95% CI [-.32, -.12].

5.3.4 Acting experience

A higher level of acting experience was associated with higher levels of altruism, $r_s = .15$, $p = .003$, 95% CI [.05, .24], print, film, and play-exposure, $r_s = .098$, 95% CI [.09, .20], $r_s = .19$, 95% CI [.09, .28], and $r_s = .25$, 95% CI [.15, .35] respectively, $df = 400$, all ps $< .05$.

5.4 Discussion

The present study took a multidimensional approach to investigating associations between fiction and empathic abilities. Consistent with the primary prediction and extending the work of previous studies examining print-exposure (e.g., Fong et al., 2013; Kidd & Castano, 2013; Mar et al., 2006, 2009), fiction-exposure via printed prose (authors), film and plays combined, and control variables, significantly predicted perspective-taking, fantasy, empathic concern and altruism. However, fiction-exposure only contributed to the predictions for altruism and fantasy (supporting Hypotheses 2 and 4). Females tend to score higher than males on the IRI (Davis, 1980), and this was the case with the present sample. There were no gender differences in altruism, or the print- or play-exposure dimensions of the FMET (similarly, Mar et al., 2006, reported no gender differences in author recognition), although males recognised more film names than females.

Individually, print-, film-, and play-exposure differentially correlated with empathic abilities. The more authors and plays participants recognised, the higher their levels of altruism. Fantasy was positively associated with all fiction media dimensions when age, gender, education and English language fluency were controlled, and with print-exposure with and without the inclusion of control variables. Effects sizes were small in magnitude, which is consistent with previous studies (Mumper & Gerrig, 2017). Specifically, the correlation between print-exposure and fantasy was similar ($r = .13$ in the present study compared to $r =$

.18 in their meta-analysis). The correlation was similar for perspective-taking ($r = .05$, non-significant, compared to .08 in the meta-analysis). In the present study, print-exposure negatively correlated with empathic concern ($r = .01$, non-significant) whereas the association was positive in Mumper and Gerrig's meta-analysis ($r = .07$), though this was in line with the negative effect reported in Dodell-Feder and Tamir's (2018) meta-analysis of causal effects ($g = -.03$). Effects on concern may be difficult to determine because, in the model of empathy used in the present research, it is mediated by empathic inferencing strategies. Future research may clarify fiction effects on the concern component, and the role of other components of empathy.

Correlations between media preferences and empathic abilities generally reflected those for media-exposure. Preference for reading novels positively correlated with fantasy and altruism and preference for plays with altruism and empathic concern. Contrastingly, preference for film negatively correlated with altruism, and TV with altruism, empathic concern and perspective-taking.[13] In order to rule out the possibility that respondents who reported a preference for film and TV represented viewers of nonfiction programmes (i.e., documentaries, "reality TV", or news), correlations between media and genre preferences were also examined. There was no relationship between the factual/documentary genre and any media preferences, and except for a negative relationship between TV and experimental/postmodern, film and TV were not associated with any specific genres. Acting experience was linked to higher levels of altruism, and with exposure to all FMET dimensions (most strongly with play-exposure). This suggests that people with experience of acting are more familiar with stories in general and tend to exhibit more helping behaviours.

---

[13] The FMET was designed to measure printed prose, live theatre, and filmed media presentations of fiction. Considering the differences in results for film and TV preferences, it would be useful to establish a TV-exposure dimension to facilitate a nuanced understanding of the impact of exposure to the range of fictional narratives available onscreen.

Enjoyment of the factual/documentary genre was associated with altruism, perspective-taking and empathic concern; however, correlations with empathic abilities tended to be stronger for the fiction genres (in line with prior research, see Mumper & Gerrig, 2017). Specifically, enjoyment of drama and romance were positively associated with perspective-taking, empathic concern and fantasy; and comedy, the most popular genre, and experimental/postmodern, the least popular genre, were both associated with higher levels of all empathic abilities measured. These findings indicate that relationships between fiction and empathic abilities are not unique to literary prose and that empathic processes may be induced, to different degrees, by a range of narrative content.

5.4.1 Fiction-engagement content and processes

Media presentation is fundamental to the way content is received (e.g., McLuhan, 1964/1994). Actors and fiction-readers, who tended to report higher levels of altruism in the present study, imaginatively construct storyworlds "from within" (Hakemulder, 2000, p. 11). In contrast, filmed media provides images of the environment, as well as visual (facial, bodily and proximal) and auditory (tonal, prosodic and musical) cues for interpreting characters' inner states and may entail a more passive engagement style. Mar et al. (2010) found that exposure to storybooks and film, but not TV, positively predicted mentalising ability in children, which aligns with the present finding that preference for TV was negatively associated with all empathic ability measures except fantasy. Preference for film negatively correlated with altruism in the present adult sample. Future research may clarify relationships between onscreen fiction forms and empathic abilities and establish any interaction with age.

Recall that Kidd and Castano (2013) proposed that active, writerly comprehension, which can strengthen readers' abilities to mentalise about others' internal states, is induced by literary texts. In this view, the stylistic features that distinguish literary from genre-fiction, such as complex and unpredictable characters (Kidd & Castano, 2018b) may stimulate

personality change and foster empathic understanding (Djikic & Oatley, 2014; Kidd & Castano, 2013, 2017a). In the present study, only comedy and experimental/postmodern genres were associated with all empathic ability measures and, along with factual/documentary, were the only genres associated with altruism. Just as certain media presentations can elicit more active engagement, these genres may entail a particularly writerly engagement style by subverting expectations and requiring multiple perspectives to be tracked: experimental/postmodern narratives incorporate techniques such as fragmentation, multiple plotlines and unreliable narration, and interpreting comedy involves the integration of incongruent mental states, verbal expressions and context.[14] However, comedy, as well as romance and drama genres, can be formulaic, and is generally associated with commercial mass media, rather than with literary media. Relationships between popular genres and empathic abilities could result from the incorporation of literary devices, or features of genre-fiction may, too, activate processes that mediate fiction-empathy relationships.

The process of becoming transported (Gerrig, 1993), or "lost" in storyworlds (Nell, 1988, p. 8), imagining story environments and simulating the experiences of fictional characters (e.g., Mar et al., 2006), appears to mediate the effects of engaging with fictional social content on attitudes (Green & Brock, 2000), empathy dimensions (Bal & Veltkamp, 2013; see also Taylor & Carlson, 1997), and prosocial behaviour (e.g. Johnson, 2012; Schellenberg, 2004). In the present framework, the IRI fantasy subscale is treated as a measure of experience sharing. However, unlike other IRI dimensions which measure self-reported empathy for real-world others, the fantasy subscale probes the trait tendency to become imaginatively absorbed in fictional stories and share in the experiences of fictional characters. In the present study this tendency was associated with increased exposure to all

---

[14] Stimuli featuring comic narratives are frequently used to test advanced mentalising abilities (e.g., Aykan & Nalçacı, 2018; Happé, 1994; Winner et al., 1998).

FMET dimensions. It is unclear how far fantasy relates to other media formats not examined in this study (previous research, for example, has reported lower levels of absorption when reading via electronic screens compared to printed literature, Mangen & Kuiken, 2014), and the extent to which it maps onto experience-sharing with real-world others.

Johnson (2012) found that differences in immersion (transportation) accounted for variation in altruistic behaviour immediately after reading; however, this finding may have been partly due to readers modelling the prosocial actions of story characters (a social learning process; Johnson, 2012; see also Gentile et al., 2009). The relationship between the FMET and self-reported altruism shown in the present study indicates that some variability in altruistic tendencies is attributable to general fiction-exposure, rather than specific to stories portraying prosocial content (characters and scenarios). Controlling for fantasy and the other IRI dimensions did not significantly alter the prediction for altruism, indicating that neither the trait tendency to become absorbed in stories, nor general empathic abilities, accounted for substantial variance. Considering the patterns involving altruism in this study, further research may explore the effects of reading on altruistic behaviours using task-based approaches. Future studies may also help to establish how far levels of immersion elicited by different media formats relate to empathic processes in the context of real-world others, and the extent to which individual differences in the fantasy trait contribute to these effects.

The observed associations among fiction media formats, genres and empathic dimensions support the view that variance in empathic abilities is partly linked to the social content of fiction, rather than to more general reading processes (Mar et al., 2006; Oatley, 1999). It may be that fiction formats which entail writerly engagement with themes and content, or those that incur high levels of immersion and the simulation of characters' inner states, support empathic abilities (see Mar et al., 2009). Alternatively, fiction effects may occur via more than one route. Prosocial motivation can arise from mentalising about others'

experiences or by sharing in those experiences vicariously (Zaki & Ochsner, 2012), and the extent to which people engage their mentalising and experience-sharing faculties during exposure to different fiction media and thematic content may vary.

Media format and thematic genre were both implicated in associations between fiction and empathic abilities, and so future studies may productively incorporate media and genre-exposure dimensions (e.g., Black et al., 2018; Kidd & Castano, 2017a) in order to establish interactions (studies controlling for fiction-exposure may also benefit from the multidimensional approach). Examining the full extent of fiction-engagement (video games, for example, may also support empathic skills, e.g., Bormann & Greitemeyer, 2015; Gentile et al., 2009) was beyond the scope of this study, but the present findings justify further research into a more comprehensive range of fiction formats, and examination of the processes underpinning their effects on prosociality.

5.4.2 Limitations and future research

The correlational design of this study does not provide evidence for causation. It could be, for example, that altruistic people are motivated to read about fictional characters, and that the complexities of comic and postmodern narratives are sought by empathic individuals (e.g., Hall & Bracken, 2011). Other variables may also be implicated, including intelligence, knowledge, verbal abilities, social needs (e.g., Djikic & Oatley, 2014; Gabriel & Young, 2011; Mol & Bus, 2011; Stanovich et al., 1995; see also Mar et al., 2009) and other personality traits (for example, dark triad traits are associated with deficits in affective, but not cognitive empathy; Wai & Tiliopoulos, 2012). Prior research has shown that fiction-exposure predicted RMET performance when gender, and traits fantasy and openness were controlled (Mar et al., 2009). The multidimensional approach of the present study showed that fiction-exposure significantly explained variance in fantasy and altruism when gender, age, education and language were controlled. This study did not address the social nature of

different modes of fiction-engagement, which may play a role. For example, appreciation of plays or comedy might reflect enjoyment of being part of a live audience. Future research may elucidate the impact of social contextual factors in relationships between fiction and empathic abilities.

A lack of correlation among empathy measures has been documented (e.g., Mar et al., 2006) and further research is required to establish how far the fiction scales used in this study would relate to performance on behavioural empathic ability measures. For example, participants may complete self-report measures in ways perceived to be socially desirable. To address this, irrelevant items were included with the empathic ability measures to mask the study's aims (average values for all IRI scales were in line with prior research, Konrath, 2013), and the FMET incorporated a points-deduction system to control for guessing. The film and play names for the FMET were selected from lists of award-winning and significant works, whereas the print scale contained both literary and popular fiction authors. Therefore, it is not possible to draw conclusions about the importance of exposure to literary fiction; however, links between empathic abilities and genre preferences (such as comedy, which is generally associated with commercial, rather than literary media), indicate that relationships between fiction-exposure and empathic abilities are not exclusive to literary works.

Examining media and genre preferences in addition to fiction media-exposure facilitated a more comprehensive perspective of fiction-engagement, though multiple tests increase the risk of false positive results. However, the observed consistencies across relationships between media-exposure and media preferences with empathic abilities suggest avenues for future research. The FMET combined scores did not significantly contribute to the predictive models for empathic concern and perspective-taking, which was surprising considering previous findings linking print-exposure with these dimensions (Mumper & Gerrig, 2017). Fifteen percent of participants were not native English speakers and may have been less

familiar with some of the names in the FMET. This could have reduced the strength of correlations which were generally weaker for print-exposure than those reported in other research (see Mumper & Gerrig's, 2017, meta-analysis). Fiction- and nonfiction-exposure tend to correlate (Mar et al., 2006, 2009), but nonfiction-exposure was not controlled in this study, and fiction authors who have also published some nonfiction works were not omitted from the FMET. Therefore, variance caused by nonfiction-exposure may have attenuated or accounted for some of the observed effects. In this study, enjoyment of the factual/documentary genre showed small, positive associations with perspective-taking, empathic concern and altruism, and it may be that person-centred nonfiction such as biography and memoir also contain the narrative aspects associated with empathic development. Future research should aim to comprehensively account for variance caused by nonfiction media-exposure.

5.4.3 Summary and moving forward

People engage with a variety of fictional genres through an increasing range of formats. This study involved the development of a new tool based on previous ARTs, which enables the measurement of exposure to fiction via print, film and plays, and associations with empathic abilities. Correlations between empathic abilities and experience of acting, as well as media and genre preferences were explored. Results revealed that fiction-exposure, via printed prose, film and theatre combined, positively predicted variance in self-reported fantasy and altruism. Media and genre preferences and acting experience differentially correlated with empathic abilities, indicating that both engagement channel and thematic content are important in relationships between fictional stories and empathy. In particular, there were significant negative correlations between preference for film and TV and some empathic abilities (film with altruism and TV with all scales). It may be that the processes

involved in engaging with film reduce empathic abilities, or a third factor may come into play, such as preference for particular content, or other motivational factors.

The findings point to a central role for media presentation in relationships between fictional social content and empathic abilities. Divergent associations between genre and media formats and empathic abilities could implicate both "content" and "process" explanations of fiction effects, which are not mutually exclusive (Mar, 2018a). Genres may provide access to particular social knowledge, and certain storylines and presentation modalities may engage specific empathic processes to different degrees. It may be that transmission of social knowledge and simulation of social experience support alternate (mentalising versus experience-sharing) routes to empathic accuracy, and the consequential activation of altruistic motivation. Further research may extend these findings and help to clarify the antecedents and consequences of fiction-engagement, and the factors that moderate effects.

The next phase of this research focused on establishing relationships between fiction-exposure and experience-sharing versus mentalising processes. In Study 1, empathic abilities were measured via self-report scales which may have been subject to socially desirable responding. Therefore, in Study 2, a behavioural measure was employed. The FMET may be extended in future studies aiming to establish a range of media dimensions of fiction- and nonfiction-exposure, but as it does not contain nonfiction dimensions, a version of the original ART was used in the next study. Study 2 aimed to examine associations between fiction-exposure and a behavioural measure of empathic accuracy, which enabled the explicit comparison of mentalising and experience-sharing processes.

Chapter 6: Relationships between Fiction-exposure and Mentalising versus Experience-sharing Strategies for Empathic Accuracy

The previous chapter identified associations between fiction-exposure and self-report empathic abilities. The present study extended this line of enquiry by assessing relationships between fiction-exposure and a behavioural measure of empathic accuracy. Empathisers can interpret a target's experience of an event by using prior knowledge and decoding facial and voice cues (mentalising), or they can use their own emotional response as a proxy for the target's (experience-sharing). An empathic accuracy task was produced, based on a previous study examining the relative efficacy of these strategies (Zhou et al., 2017). It enabled assessment of the effect of inferencing strategy on accuracy for targets' emotional ratings for positive, neutral and negative images. The main hypothesis that fiction-exposure would positively predict empathic accuracy, and that this effect would be stronger for participants that used the experience-sharing strategy, was tested. The two strategies entailed similar levels of error but in opposite directions, highlighting nuanced differences between empathic inferencing strategies. Empathic accuracy varied as a function of target and valence and was positively predicted by lifetime fiction-exposure.

6.1 Introduction

Empathic accuracy, the ability to accurately infer the internal experiences of another person or target (Ickes, 1997), represents an essential component of intersubjectivity. It enables behaviour prediction (Nichols & Stich, 2003), and is associated with positive relationships (Castano, 2012), interpersonal cooperation and prosocial behaviour (Paal & Bereczkei, 2007). Zaki and Ochsner (2012) modelled two pathways to empathy: mentalising (explicitly interpreting verbal and nonverbal cues) and experience-sharing (vicariously sharing in the target's experience). This model was developed to incorporate the empathic

accuracy component (the measurable ability to accurately infer a target's mental state content), and this was outlined in Chapter 2.

Zhou et al. (2017) compared the efficacy of two strategies for empathic accuracy. Their two critical empathic inferencing conditions can be mapped onto the empathy model defined in Chapter 2, in order to facilitate a comparison of mentalising and experience-sharing approaches. In their study (Experiments 1-4), participants were asked to estimate the emotional ratings that target individuals ("experiencers") had previously given in response to a range of positive, negative and neutral photographs. Participants provided their estimations after viewing short videos of the experiencers' facial expressions ("theorisation" condition), or the photographs themselves ("simulation" condition). Therefore, participants in the theorisation condition engaged in a mentalising task (which required them to infer affective content), whereas participants in the simulation condition engaged in an experience-sharing task, where they were able to use their own reactions to the pictures as proxy for the experiencers'. Participants in a third "simultaneous" condition viewed both the photographs and the videos simultaneously. In Experiments 2-3, participants were further split into a "bound" or "free choice" condition. In the former they were assigned to their condition, whereas in the latter they were invited to select their preferred (theorisation or simulation) condition following video training on each method.

Using the simulation strategy led to higher empathic accuracy compared to the theorisation strategy and participants tended to overestimate the insight gained through theorisation compared to simulation. Indeed, when financially incentivised to perform well—participants were informed that they would receive additional payment if their performance reached the 80th percentile (Experiments 2-4)—participants in the free-choice condition tended to self-select into the less effective theorisation (mentalising) group. Not only were the two strategies for interpreting mental states shown to be unequal, participants also misjudged

their relative utility. Zhou et al. (2017) argued that their finding that participants were reluctant to use their own experience as a guide for estimating someone else's was analogous to findings from the field of affect forecasting where participants tend to under-appreciate the value of using another person's experience as a guide for their own (Gilbert, Killingsworth, Eyre & Wilson, 2009).

6.1.1 Fiction-exposure and empathic strategies

The correlational research reviewed in Chapter 3 indicated that lifetime exposure to fiction is positively associated with empathic accuracy via both experience-sharing and mentalising routes. When measures were organised into experience-sharing and mentalising (as well as concern and prosocial behaviour) categories, relationships with experience-sharing were shown to be the strongest. However, experience-sharing was measured using dispositional scales rather than behavioural tasks, whereas mentalising was measured using a combination of both. The present study aimed to assess the relationship between fiction-exposure and both strategies for empathic accuracy using a behavioural measure.

Frequent fiction-readers are regularly exposed to the emotional states of characters where the range of circumstances and experiences may be very different to the reader's own. If readers become practiced at using their own mental apparatus to make sense of characters' experiences, they may also tend to use their own beliefs, motivations and emotions as proxy for real-world others. This could result in an improved ability to recognise mental states in others based on their own emotional experiences, and tendency to trust this capability. Zhou et al.'s (2017) results suggest that experience-sharing is a beneficial strategy for accurately interpreting the emotional states of targets. Fiction-exposure may positively predict both the ability to accurately interpret emotions using this strategy, and the willingness to engage it.

6.1.2 Other factors affecting empathic accuracy

Zhou et al. (2017) acknowledged that "simulation" may not represent the most effective strategy across all contexts, and that situational factors would likely moderate the value of each strategy. The present study aimed to account for variance in empathic accuracy that could be caused by two contextual variables inherent to the stimulus materials: identity of the target experiencer and valence of the target emotion.

6.1.2.1 *Target.* In their two-systems account, Zaki and Ochsner (2012) referred to research indicating that people are less likely to adopt an experience-sharing approach with targets that they perceive as dissimilar to themselves (Hein et al., 2010). Zhou et al. (2017) suggested that participants' over-confidence in their ability to infer mental states from facial expressions may be due to the tendency to overestimate dissimilarity between self and others and so underestimate their tendency to have a similar emotional response to a target. This can lead to an ingroup advantage in accuracy (Adams et al., 2010; Matsumoto et al., 2009). Therefore, the identity of the target individual may impact the extent to which experience-sharing and mentalising strategies are successfully engaged (e.g., Barrett, Mesquita & Gendron, 2011).

6.1.2.2 *Valence.* People show attentional bias toward emotionally salient information ("affect-biased attention"; Todd, Cunningham, Anderson & Thompson, 2012) when viewing complex scenes (Humphrey, Underwood & Lambert 2012). Positive and negative emotion expressions are associated with differences in the neural activity of the perceiver (Kilts, Egan, Gideon, Ely & Hoffman, 2003) and are recognised at different speeds (positive expressions tend to be recognised faster than negative expressions, e.g., Leppänen & Hietanen, 2003). Little research has examined where variations in accuracy occur between emotion expressions of differing valence in neurologically typical adults, though studies have indicated interactions with psychiatric disorders (e.g., borderline personality disorder; Unoka, Fogd,

Füzy & Csukly, 2011) and that both depressive patients and controls show lower accuracy for sad expressions compared to happy ones (Surguladze, Young, Senior, Brébion, Travis & Phillips, 2004). The present study aimed to examine how far participants' estimates for the mental states of targets were accurate, the direction of any error (i.e., whether errors tend to be over-positive or over-negative) and the extent to which this would vary as a function of valence.

6.2 The Present Study

Considering the potential impact of valence and target, the present study diverged from Zhou et al.'s (2017) procedure in two ways: first, stimuli were produced containing six targets. This was aimed at increasing generalisability and to enable differences in identification with targets to be accounted for in the analysis (Zhou et al.'s participants were assigned to one of twelve targets and this source of variance was not addressed in their analysis). Second, overall differences in empathic accuracy were computed using absolute values in line with Zhou et al., and raw scores in order to establish differences in the direction of errors in empathic accuracy. Whereas converting scores on the dependent variable into absolute values would result in a compression of the variance associated with the full range of positive and negative responses, using the full scale via the computation of raw scores would facilitate a more nuanced measure of empathic accuracy. The aims of this study were to: replicate Zhou et al.'s finding that the experience-sharing strategy leads to higher empathic accuracy compared to mentalising (Hypothesis 5 of the hypotheses articulated in Chapter 4), examine the direction of errors across levels of valence and target (Hypotheses 6-8), predict empathic accuracy from lifetime fiction-exposure (Hypothesis 9) and examine the moderating role of strategy (Hypothesis 10), predict empathic strategy choice from lifetime fiction-exposure (Hypothesis 11) and explore the impact of identification with the target.

6.3 Method

6.3.1 Participants

Zhou et al. (2017) based their sample size on the simple heuristic that each "experiencer" (they had 12) was paired with at least two participants in each condition. This approach yielded a large effect of condition, $t(70) = 7.26$, $p < .001$; they reported a common language effect size (CL)[15] of 92.3% for the difference between theorisation and simulation with 24-25 participants per group (Experiment 1, total $n = 73$). Further, Mumper and Gerrig's (2017) meta-analysis correlating fiction-exposure with RMET (93%) or actor-intention vignette scores (7%), $r = .21$, $p < .001$, $d = .43$, two-tailed, $\alpha = .05$, indicated that a sample size of 200 would be required to detect the correlation between fiction-exposure and empathic accuracy task performance at > 80% power ($N = 200$ for .86 power). Power analyses were conducted using SPSS syntax (Appendix C), and 204 participants were recruited via the Qualtrics platform in return for financial compensation. However, inspection of the data revealed that the sample consisted entirely of females due to a fault in the online screening logic and so a further 120 male participants were recruited (total $N = 324$). Four participants were excluded due to reporting a technical (stimulus loading) error, three participants were excluded for straight-lining (selecting the same response throughout the study), and 15 participants were excluded for response durations greater than three standard deviations from the mean, resulting in a final sample size of 302. The sample consisted of 100 males (33%) and 202 females (67%) aged 18-86 ($M = 50.2$, $SD = 15.8$), resident in the UK (34%), USA (37%) or Canada (29%).

---

[15] The CL statistic (McGraw & Wong, 1992) reflects the probability that a score sampled at random from one distribution would be greater than that of another distribution. The value here indicates that a score selected from the simulation condition would be higher than a score selected from the theorisation condition in 92.3 out of 100 cases.

6.3.2 Materials

6.3.2.1 Development of empathic accuracy test stimuli. It was not possible to access Zhou et al.'s (2017) original materials, and so the empathic accuracy stimuli were developed from scratch. In line with Zhou et al., a group of students (five females, three males, aged 23-45, $M = 31.83$, $SD = 9.28$) were recruited to act as target "experiencers" voluntarily or in return for course credit if applicable. They were informed that their data would be used as stimuli in a set of future experiments that aimed to "explore how adults can improve their abilities to understand what other people are thinking and feeling". The experiencers were asked to rate their emotional responses to 60 pictures from the Geneva Affective Picture Database (GAPED; Dan-Glauser & Scherer, 2011) on a 9-point Likert-style scale from "extremely negative" to "extremely positive" (the scale used the same labels as Zhou et al., 2017). A variety of positive, negative and neutral pictures were selected (20 each), in order to generate a range of emotional responses. All materials were administered in Qualtrics and participants took part on a computer in a lab, with the experimenter present. The images were presented in a random order and each appeared for 7 seconds. With the participants' permission, their faces were filmed through a webcam for the duration of the task. Participants were given the following instruction:

> During this exercise, you will be filmed through the webcam. Try to sit fairly still with your hands away from your face. You don't need to worry about the camera as we will only be using small sections of the footage. […] We are interested in your emotional reactions to the image, so don't try to hide how you are feeling—you can let your facial expressions reflect your true feelings about the image.

During participation, the computer screen was simultaneously recorded. This enabled the two video streams to be aligned using video editing software (Final Cut Pro) and edited into 60 5-second clips of the participants' facial reactions to each picture. The 5-second clips

were extracted from the point at which the image had loaded (the additional two seconds accounted for issues with the loading of images).

The video clips presented alongside a still image of the experiencer formed the stimuli for the theorisation condition, and the GAPED pictures presented alongside the same still image of the experiencer formed the stimuli for the simulation condition (Figure 5). These were neutral screenshots taken prior to each experiencer's exposure to the images, and they remained the same across each trial for the experiencer. The still image of the experiencer was included in both conditions so that participants were equally aware of the experiencers' demographic characteristics. The materials were presented in Qualtrics. The final stimulus set consisted of 60 trials across six experiencers (four females and two males, aged 19-45), with 10 trials per experiencer, plus two "practice" experiencers (one male aged 40 and one female aged 65) with five trials in total (two and three per experiencer, respectively) in both the theorisation and simulation conditions. The practice experiencers were selected due to technological problems resulting in incomplete datasets for those experiencers. The distribution of positive, negative and neutral images across the six experiences is shown in Figure 6.
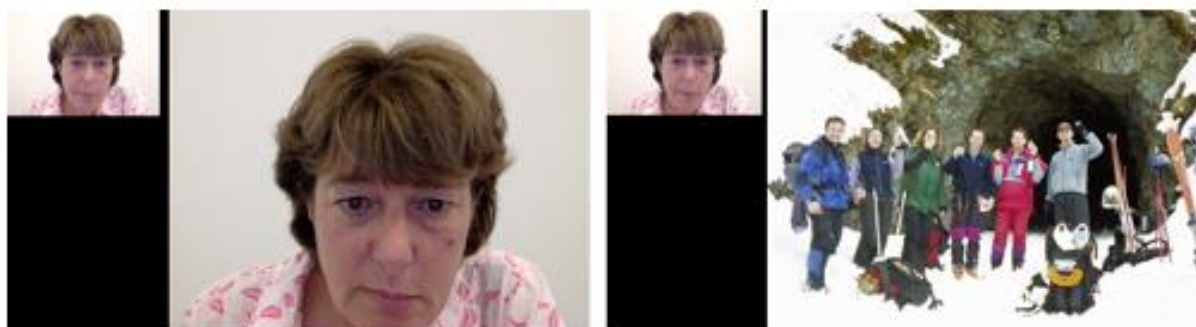
Figure 5. Screenshot of the theorisation (mentalising) condition practice stimuli (left panel) and the simulation (experience-sharing) condition practice stimuli (right panel).

6.3.2.2 Empathic accuracy test. Participants were asked to predict the emotional ratings of the six target experiencers in response to the 60 pictures from the GAPED database that the

experiencers viewed. Strategy (theorisation vs. simulation) was varied between participants: participants were randomly assigned to either the simulation (experience-sharing) condition, in which they viewed the same pictures as the experiencers, or the theorisation (mentalising) condition, in which they viewed video footage of the experiencers reacting to the pictures. Participants in both conditions were given the following instruction:

> In a previous study, we asked the experiencers to look at some pictures. We asked them to tell us how they felt about each picture on a scale from "extremely negative" to "extremely positive". We call this their "emotional rating". Your task is to estimate the experiencer's emotional rating for each picture that they viewed, using the same scale.

Trials in each condition lasted for 5 seconds. After each trial, participants were asked two questions: first, either, "what was your emotional rating for this clip of the experiencer?" (theorisation participants) or "what was your emotional rating for this picture?" (simulation participants). These questions were designed to address the possibility that the participants' own feelings in response to the video clips would interfere with their rating for the target in the theorisation condition. Second, in both conditions, participants were asked, "what was the experiencer's emotional rating for this picture?" The differences between responses to this question and the experiencers' actual ratings were used to measure empathic accuracy (calculated by subtracting the experiencer's original rating from the participant's estimate and so higher difference scores represented lower empathic accuracy). To both questions, participants responded using the same scale that the experiencers had used. Internal consistency was good ($\alpha = .87$).

6.3.2.3 Empathic strategy choice. Following completion of the empathic accuracy test, participants were informed that there were two versions of the task and presented with examples of each via looped videos (taken from the practice experiencer sets). They were then asked the question, "which option you think would be the most successful in enabling you

to estimate the experiencers' ratings as accurately as possible?" Strategy choice was recorded via a single dichotomous "theorisation" versus "simulation" item.

6.3.2.4 Identification with target. An ingroup advantage (Matsumoto et al., 2009) could impact the effect of experiencer. Before each empathic accuracy experiencer block, participants viewed a still image of the experiencer and were asked to indicate how similar they considered the experiencer to be to themselves on a single-item sliding scale from "we are completely different" (0) to "we are completely alike" (100).
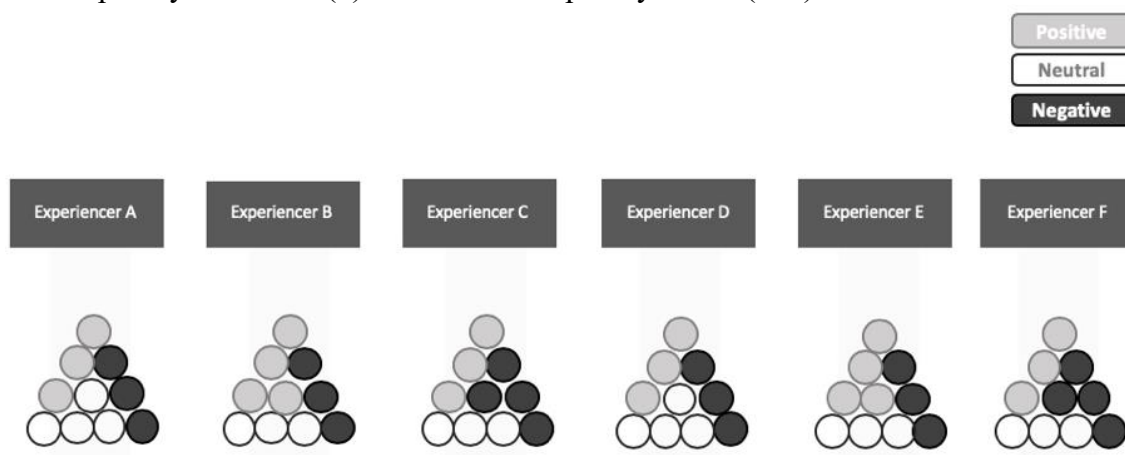
Figure 6. Distribution of positive, negative and neutral images across the six experiencers.

6.3.2.5 Fiction-exposure. This study used Mar et al.'s (2006) revised version of the ART (ART-R), which incorporates ten genres (five fiction, e.g., romance, thriller; and five nonfiction, e.g., science, business) across mutually exclusive fiction and nonfiction categories. The inclusion of the nonfiction dimension allows nonfiction-exposure to be controlled. The original ART-R contains 50 fiction authors, 50 nonfiction authors and 40 foils. The test was updated for the current study: one original name from the nonfiction category, "Norman Mailer," was replaced with "Mary Beard", due to the author's profile as an author of both fiction and of nonfiction (Mailer won the Pulitzer Prize in both categories), and "M. D. Johnson Spencer" was amended to "Spencer Johnson". A further five names (10%) were added to each of the fiction and nonfiction categories and evenly spread across genres to ensure that recent works were incorporated (each of the additions had been published or

republished since 2013). Foil-checking tends to be low (Mar et al., 2006) and so the foils

section was left unedited. The updated version of the ART-R used for this study contained 55

fiction names, 55 nonfiction names, and 40 foils (150 items total; Appendix D). A reliability

analysis indicated that internal consistency was very good (α = .94; N = 139),[16] and item-

deletion did not improve scale reliability in any case.

6.3.2.6 Perspective-taking. Self-report data on trait perspective-taking were collected

using the 7-item perspective-taking subscale of the IRI (described in Chapter 2): perspective-

taking measures the dispositional tendency to mentalise about other people's points of view

(e.g., "I try to look at everybody's side of a disagreement before I make a decision"). The

scale showed good reliability with this sample (α = .77).

6.3.3 Procedure

After providing consent and demographic information, participants completed the ART-

R, perspective-taking scale, and empathic accuracy test (in their randomly assigned condition)

and answered the empathic strategy question (Figure 7). With the exception of the empathic

strategy measure, which always followed the empathic accuracy test, the order of tasks was

randomised, and within-task items were internally randomised (in the empathic accuracy test,

trials were randomised within experiencer blocks, which were randomised within conditions).

6.3.4 Data analysis

Ninety-five percent confidence intervals were bias-corrected and accelerated using

bootstrapping (N = 1000).[17] Empathic accuracy scores were calculated by subtracting the

target experiencer's actual rating from the participant's rating for the experiencer, for each

trial, averaged across the 60 trials. Therefore, a value of zero indicates no difference between

---

[16] The following 11 items received no hits and therefore accounted for zero variance and were not included in the reliability analysis: Douglas Rushkoff, Smir D. Aczel, Stephen R. Cowley, Barry Z. Posner, Margarita Azmitia, Oscar Barbarin, Reuben Baron, Carl Corter, Frank Keil, Lynn Liben and Frances Fincham.
[17] This increased the speed of the procedure compared to the 5000 bootstrap replicates used in the previous analysis, and was above the generally sufficient number of 500 (Pattengale, Alipour, Bininda-Emonds, Moret & Stamatakis, 2009).

participants' estimates and experiencers' actual ratings, values closer to zero indicate higher

empathic accuracy and larger values indicate lower empathic accuracy. Absolute difference

values provide the magnitude of errors (differences between participant ratings for

experiencers and experiencers' own ratings) in line with Zhou et al.'s (2017) approach, and

raw difference scores establish the general direction of errors. Raw negative values indicate

that participants' average estimates were over-negative whereas raw positive values indicate

that the participants' average estimates were over-positive.
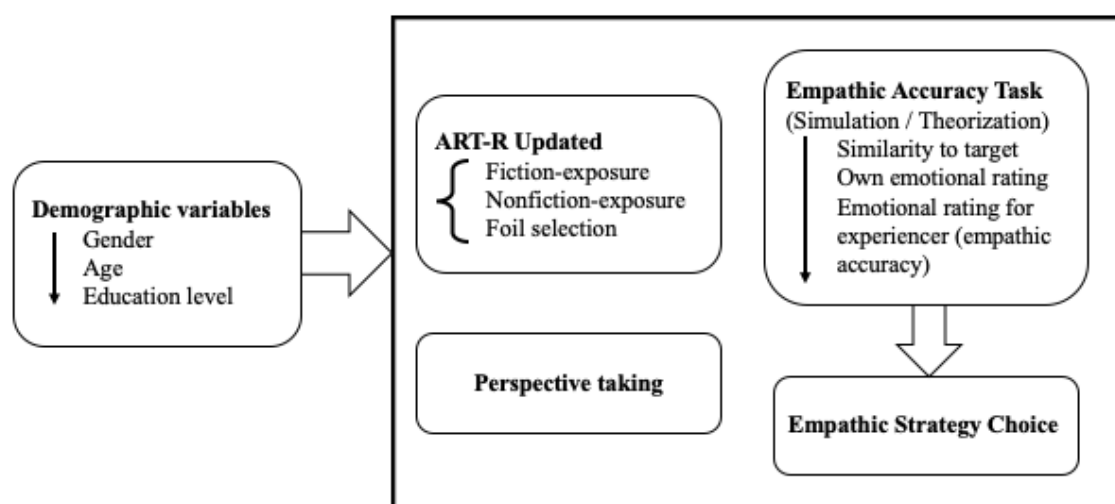
Figure 7. Experimental procedure. After demographic measures, the order of tasks was randomised
(with the exception of empathic strategy choice which always followed the empathic accuracy task).
Items were randomised within measures.

6.4 Results

6.4.1 Preliminary analyses

6.4.1.1 Scale scores. Average ART-R, perspective-taking and absolute and raw

difference scores are presented in Table 7. Participants recognised more fiction authors than

nonfiction authors in both conditions (in line with the ART-R scores reported by Mar et al.,

2006, although both means were lower than those reported in Mar et al.'s study), and foil-

selection was low. The theorisation group reported marginally higher perspective-taking

tendencies compared to the simulation group, $p = .047$, but after correcting the error rate for

the testing of both absolute and raw scores, this effect is non-significant. Absolute difference

scores indicated that participant errors were broadly equal when using the simulation compared to the theorisation strategy, $p = .118$. Raw difference scores demonstrated a more nuanced pattern of results with participant estimates in the theorisation condition tending to be over-negative, whereas estimates in the simulation condition tended to be over-positive.

6.4.1.2 Inter-scale associations. A robust (Spearman) correlation analysis was conducted to establish relationships among the demographic variables, the ART-R, perspective-taking and outcome variables, organised by condition (Tables 8a, 8b). Fiction-exposure, nonfiction-exposure and foil selection correlated in both conditions. Gender and perspective-taking were associated in the simulation condition, with females ($M = 17.24$, $SD = 5.57$) scoring higher than males ($M = 14.60$, $SD = 4.24$), $t(132.26) = -3.28$, $p = .001$, 95% CI [-4.22, -1.04]. Education level was positively associated with fiction-exposure in the theorisation condition, and age positively correlated with fiction-exposure in both conditions and with nonfiction-exposure in the theorisation condition.

Raw difference scores showed a statistically significant association with only one variable, fiction-exposure, and only in the theorisation condition. This indicated that empathic accuracy based on video footage of the experiencer increased (because higher raw difference scores indicate lower accuracy) when participants had been exposed to more fiction (although this could represent a spurious correlation). The null finding in the simulation condition may have been due to the positive and negative signs of raw scores cancelling out across aggregate means. The absolute scores, which provide a more accurate measure of the magnitude of mean differences, correlated with fiction-exposure and nonfiction-exposure in both conditions, such that higher levels of accuracy (lower difference scores) were associated with higher levels of both fiction- and nonfiction-exposure. These effects were stronger for fiction-exposure. Correlations between fiction-exposure and raw and absolute difference scores are presented in Figure 8.

Table 7

Means (Standard Deviations) and 95% Confidence Intervals for Overall Scale Scores and for

Each Strategy Group.

Measure	Grand Mean	Simulation	Theorisation	t	p
Fiction-exposure	9.38 (8.55) [8.46, 10.37]	8.69 (8.06) [7.47, 9.97]	10.11 (9.00) [8.73, 11.48]	-1.44	.15
Nonfiction-exposure	3.45 (4.06) [2.03, 3.93]	3.25 (3.97) [2.65, 3.93]	3.67 (4.14) [3.01, 4.39]	-.90	.37
Foil selection	0.41 (0.92) [0.31, 0.52]	0.35 (0.76) [0.25, 0.48]	0.47 (1.07) [0.31, 0.67]	-1.08	.29
Perspective-taking	16.90 (5.07) [16.31, 17.51]	16.34 (5.29) [15.57, 17.16]	17.50 (4.77) [16.71, 18.35]	-2.0	.05
Absolute difference scores	1.45 (0.32) [1.42, 1.49]	1.42 (0.32) [1.37, 1.47]	1.48 (0.32) [1.43, 1.53]	-1.57	.12
Raw difference scores	0.01 (0.48) [-0.05, 0.06]	0.12 (0.44) [0.05, 0.19]	-0.11 (0.50) [-0.19, -0.02]	4.13	< .001

Note. 95% confidence intervals (bias-corrected and accelerated using bootstrapping $N = 1000$) are presented in brackets.
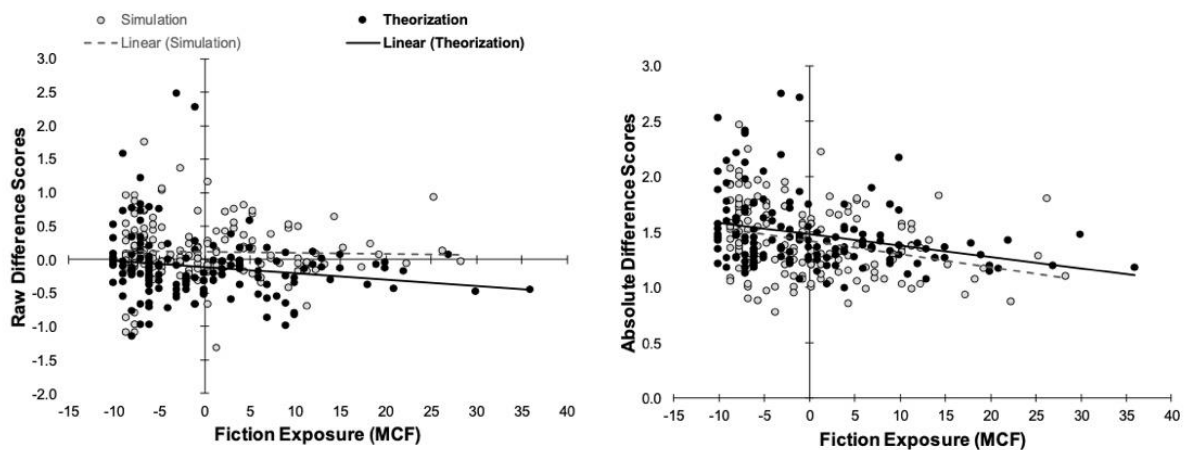
Figure 8. Correlations between fiction-exposure (mean centred form) and raw (left) and absolute (right) difference scores.

Table 8a

Correlations Among Variables within Simulation Condition.

	Gender	Education	FE	NE	Foils	PT	Raw	Abs	Strategy choice
Simulation condition (*N* = 155)									
Age	-.06 [-.23, .12]	.01 [-.16, .17]	.30 *** [.16, .45]	.13 [-.03, .08]	.03 [-.13, .19]	-.03 [-.18, .14]	-.01 [-.16, .14]	-.05 [-.20, .11]	.06 [-.11, .002]
Gender	-	-.10 [-.25, .05]	.12 [-.04, .27]	-.05 [-.21, .11]	.03 [-.13, .19]	.25 ** [.11, .40]	-.03 [-.19, .13]	-.02 [-.20, .15]	-.01 [-.18, .14]
Education		-	.02 [-.14, .18]	.12 [-.03, .27]	-.01 [-.17, .15]	.02 [-.15, .18]	-.001 [-.17, .17]	-.14 [-.31, .03]	.04 [-.11, .18]
Fiction-exposure (FE)			-	.61 *** [.49, .71]	.28 *** [.13, .41]	.05 [-.09, .21]	-.08 [-.23, .08]	-.30 *** [-.41, -.17]	-.08 [-.24, .10]
Nonfiction-exposure (NE)				-	.29 *** [.14, .42]	.06 [-.10, .23]	-.01 [-.18, .15]	-.21 ** [-.32, -.09]	-.08 [-.23, .08]
Foil selection (Foils)					-	-.06 [-.21, .08]	.05 [-.13, .21]	.02 [-.13, .16]	-.06 [-.20, .09]
Perspective-taking (PT)						-	.16 [.01, .31]	-.09 [-.25, .09]	-.05 [-.22, .11]
Raw difference (Raw)							-	.08 [-.17, .35]	-.01 [-.15, .15]
Absolute difference (Abs)								-	-.001 [-.17, .16]

Note. * *p* < 05, ** *p* < .01, *** *p* < .001, 95% bias-corrected and accelerated confidence intervals (bootstrap *N* = 1000) are presented in brackets.

Table 8b

Correlations Among Variables within Theorisation Condition.

	Gender	Education	FE	NE	Foils	PT	Raw	Abs	Strategy choice
Theorisation condition (N = 145)									
Age	-.14 [-.28, .02]	.15 [-.01, .29]	.36 *** [.22, .50]	.29 *** [.14, .43]	.06 [-.11, .22]	-.08 [-.24, .23]	-.06 [-.23, .10]	-.004 [-.18, .17]	.08 [-.08, .24]
Gender	-	.02 [-.14, .18]	.12 [-.04, .29]	.02 [-.14, .18]	-.02 [-.17, .14]	.15 [-.01, .33]	-.01 [-.16, .14]	-.12 [-.28, .04]	.09 [-.07, .26]
Education		-	.17 * [.01, .32]	.22 ** [.04, .38]	.17 * [-.02, .34]	.12 [-.05, .28]	-.11 [-.28, .05]	-.20 * [-.35, -.04]	-.06 [-.24, .11]
Fiction-exposure (FE)			-	.67 *** [.56, .75]	.20 * [.04, .36]	.10 [-.07, .25]	-.20 * [-.36, -.05]	-.29 ** [-.39, -.19]	.01 [-.17, .17]
Nonfiction-exposure (NE				-	.24 ** [.07, .39]	.10 [-.07, .26]	-.09 [-.25, .07]	-.18 * [-.29, -.07]	-.01 [-.17, .16]
Foil selection (Foils)					-	.13 [-.03, .28]	-.01 [-.18, .15]	.02 [-.15, .25]	.08 [-.09, .24]
Perspective-taking (PT)						-	.07 [-.10, .25]	-.03 [-.20, .15]	-.08 [-.24, .07]
Raw difference (Raw)							-	.31 *** [-.01, .54]	-03 [-.18, .16]
Absolute difference (Abs)								-	-.10 [-.25, .07]

Note. * *p* < 05, ** *p* < .01, *** *p* < .001, 95% bias-corrected and accelerated confidence intervals (bootstrap *N* = 1000) are presented in brackets.

6.4.1.3 Empathic accuracy as a function of participant gender. Nether raw nor absolute scores differed as a function of participant gender in either the simulation or theorisation conditions, all *p*s > .05. Inclusion of the perspective-taking covariate did not alter this result.

6.4.1.4 Identification with target.

There were no significant associations between identification with target and raw differences between ratings for self and experiencer, with the exception of Experiencer A in the theorisation condition, *r*(145) = .22, *p* = .008, 95% CI [.05, .38]. Nor was identification with target associated with empathic accuracy in either condition. Therefore, this variable was not included in further analyses.

6.4.2 Primary Analyses

6.4.2.1 Empathic accuracy as a function of empathic strategy. Independent *t*-tests using the absolute and raw difference scores addressed the hypothesis that there would be an effect of strategy on empathic accuracy. As shown in Table 7, using absolute difference scores yielded a non-significant effect of condition, indicating that the average error was similar in magnitude across the two conditions. However, the effect on raw difference scores was statistically significant, *t*(300) = 4.13, *p* < .001, 95% CI [.12, .33], such that simulation condition estimates tended to be over positive (*M* = .12, *SD* = .44), whereas theorisation condition estimates tended to be over-negative (*M* = -.11, *SD* = .50). While this represents a very small mean difference in the context of the 9-point response scale, the size of the difference would have been compressed as a result of positive and negative values cancelling out across the dataset. The assessment of the direction of error using the raw scores supported the prediction that accuracy would differ between strategies, but it did not replicate Zhou et al.'s (2017) finding that accuracy was greater in the simulation condition.

6.4.2.2 Empathic accuracy as a function of strategy, experiencer and valence. Figure 9 shows the overall pattern of results for strategy, experiencer and valence and illustrates the different information provided by the absolute and raw difference scores. Both panels show that accuracy tended to be higher for neutral images (as differences scores were closer to 0). The raw scores revealed that participants tended to be over-negative when estimating experiencer ratings for positive images, and over-positive when estimating experiencer ratings for negative images. In other words, participants' estimates for the experiencers were generally more conservative than the experiencers' own ratings. The effects of strategy, target and valence were analysed using two 2(strategy: simulation, theorisation) × 6(experiencer: A, B, C, D, E, F) × 3(valence: positive, neutral, negative) mixed analyses of variance on the absolute and raw difference scores.

6.4.2.2.1 Assumption checks. There were several outliers and the assumption of homogeneity of variances was violated in several conditions, as assessed using Levene's test. However, ANOVA should be fairly robust to deviations from normality with a large sample size, and to heterogeneity with fairly equal sample sizes (Norman, 2010), and so the analysis proceeded as planned. The assumption of sphericity was violated, indicated by Mauchly's test (which may have been due to over-sensitivity with large samples; Weinfurt, 2000), and so Greenhouse-Geisser corrected values are reported (Maxwell & Delaney, 2004).

6.4.2.2.2 Mixed ANOVAs. The same pattern of results was found using both sets of scores: using the absolute scores revealed significant main effects of strategy, $F(1, 300) = 4.42$, $p = .036$, $\eta_{p2} = .015$, valence, $F(2, 600) = 375.79$, $p < .001$, $\eta_{p2} = .56$, and experiencer, $F(4.22, 1265) = 131.79$, $p < .001$, $\eta_{p2} = .31$, and a significant three-way interaction effect, $F(7.62, 2286) = 39.60$, $p < .001$, $\eta_{p2} = .12$. In light of the three-way interaction, the data were then split by condition to assess the simple two-way interactions. The assumption of sphericity was violated for both groups (Mauchly's test $p < .001$) and so Greenhouse-Geisser

adjusted values are reported. The interaction between experiencer and valence was statistically significant in both the simulation, $F(7.08, 1089) = 12.34$, $p < .001$, $\eta_{p2} = .07$ and theorisation, $F(6.60, 962) = 98.02$, $p < .001$, $\eta_{p2} = .40$, conditions. Using the raw scores also showed significant main effects of strategy, $F(1, 300) = 21.08$, $p < .001$, $\eta_{p2} = .07$, valence, $F(1.27, 379) = 745.49$, $p < .001$, $\eta_{p2} = .71$, and experiencer, $F(4.29, 1286) = 38.18$, $p < .001$, $\eta_{p2} = .11$, and a three-way interaction effect, $F(7.58, 2272) = 34.10$, $p < .001$, $\eta_{p2} = .102$. Simple effects analyses revealed significant interactions between experiencer and valence in both the simulation condition, $F(6.29, 968) = 116.12$, $p < .001$, $\eta_{p2} = .43$, and the theorisation conditions, $F(8.49, 1240) = 317$, $p < .001$, $\eta_{p2} = .69$. Independent t-tests were conducted to assess whether there were differences between the simulation and theorisation conditions for experiencer, valence, and experiencer $\times$ valence interactions for both the absolute (Appendix E.1) and raw difference scores (Appendix E.2).

6.4.2.3 Fiction-exposure as predictor of empathic accuracy. A multiple linear regression was conducted to predict empathic accuracy using absolute difference scores from fiction-exposure, nonfiction-exposure and foil selection on the ART-R. All assumptions were met (linearity, normality, homoscedasticity, independence of residuals, and lack of multicollinearity). Three outliers were retained in the analysis (there were no extreme leverage points). The ART-R dimensions predicted 9% variation in absolute difference scores, $F(3, 298) = 9.32$, $p < .001$, $R^2 = .09$, adj. $R^2 = .08$. Only fiction-exposure, $\beta = -.28$, $p < .001$, 95% CI [-.02, -.005], significantly contributed to the prediction, such that for each fiction author recognised, empathic accuracy scores increased by .011 ($B = -.011$).[18] The same procedure using the raw difference scores revealed similar results. There was independence of residuals as indicated by a Durbin-Watson statistic of 2.08. There was

---

[18] Negative beta values represent increases rather than a decreases because empathic accuracy was computed as inverted difference scores.

linearity and homoscedasticity as assessed by visual inspection of partial plots and studentised residuals versus unstandardised predicted values, respectively. The data were normally distributed as assessed via visual inspection of the histogram and P-P plot. All correlation coefficients were < .7 and Tolerance values were > 0.1 (the lowest was .47 indicating no issues with multicollinearity). Four outliers were identified, but as only one leverage value was above 0.2 (0.25) and no Cook's distance values were above 1 (highest value = 0.16), none were excluded. The model predicted 4% of variance in raw difference scores, $R^2 = .19$, adj. $R^2 = .035$, $F(3, 298) = 3.57$, $p = .014$. Fiction-exposure, $\beta = -.17$, $p = .031$, 95% CI [-.02, -.001], and foil selection $\beta = .14$, $p = .027$, 95% CI [.008, .13], significantly contributed to the model. Inclusion of age, which was positively associated with fiction-exposure in the present sample, revealed the same pattern of results, and age did not significantly contribute to either model.

A moderator analysis was conducted to assess the role of condition in the relationship between fiction-exposure and empathic accuracy. Mean centring was not used in the model as values of "0" were meaningful. The interaction term did not produce a significant increase in explained variance in the absolute, $F(1, 298), = .15$, $p = .70$, or raw scores $F(1, 298) = 1.50$, $p = .23$, indicating that the positive relationship between fiction-exposure and empathic accuracy did not differ across the simulation versus theorisation conditions. These results supported the hypothesis that fiction-exposure would positively predict empathic accuracy, but the effect was not shown to vary as a function of condition.

6.4.2.4 Fiction-exposure and strategy condition as predictors of empathic strategy choice. In line with Zhou et al.'s (2017) findings, participants indicated that viewing videos of the experiencers (theorisation, 64%) would represent the more effective strategy for estimating their emotion ratings in response to pictures, compared to viewing the same pictures (simulation, 36%). Of participants who had completed the task in the theorisation

condition, 48% selected theorisation as the most effective strategy and 52% selected

simulation. In contrast, in the simulation condition 80% of participants indicated that

theorisation would represent the most effective strategy, compared to 20% that selected

simulation. A Chi-square test (with all expected cell counts > 5) revealed that the association

between previous strategy condition and strategy choice was statistically significant, $\chi^2(1) =$

34.30, $p < .001$.

Binary logistic regression was used to predict empathic strategy choice from fiction-

exposure, nonfiction-exposure, foil-selection and strategy condition. Relationships between

the continuous variables and the logit of the outcome variable were linear, as assessed using

the Box-Tidwell procedure (Box & Tidwell, 1962) with Bonferroni correction (alpha was

divided by the eight interaction terms, resulting in statistical significance accepted at $p <$

.00625; Tabachnick & Fidell, 2007), all ps > .016. There were no significant outliers. The

model was statistically significant $\chi^2(4) = 36.25$, $p < .001$ (the Hosmer and Lemeshow test

indicated that it was a good fit for the data, $\chi^2[8] = 10.49$, $p = .23$) and explained 16%

(Nagelkerke R^2) of the variance in strategy choice. Sensitivity was 72.5%, specificity was

54.2%, predictive value for theorisation was 74% and for simulation was 52.3%, with the

model correctly classifying 66% of all cases. Results revealed that fiction-exposure did not

predict strategy choice: only the condition in which participants had completed the empathic

accuracy task significantly contributed to the prediction, $B = -1.50$, $p = .001$, 95% CI [-2.09, -

1.00]. Participants assigned to the theorisation condition for the task were 18% less likely to

select the theorisation strategy compared to participants assigned to the simulation condition

(calculation: odds ratio = .22 / [1 + odds ratio] = 0.18). Participants tended to consider the

condition they had completed the task in as less effective for estimating others' emotional
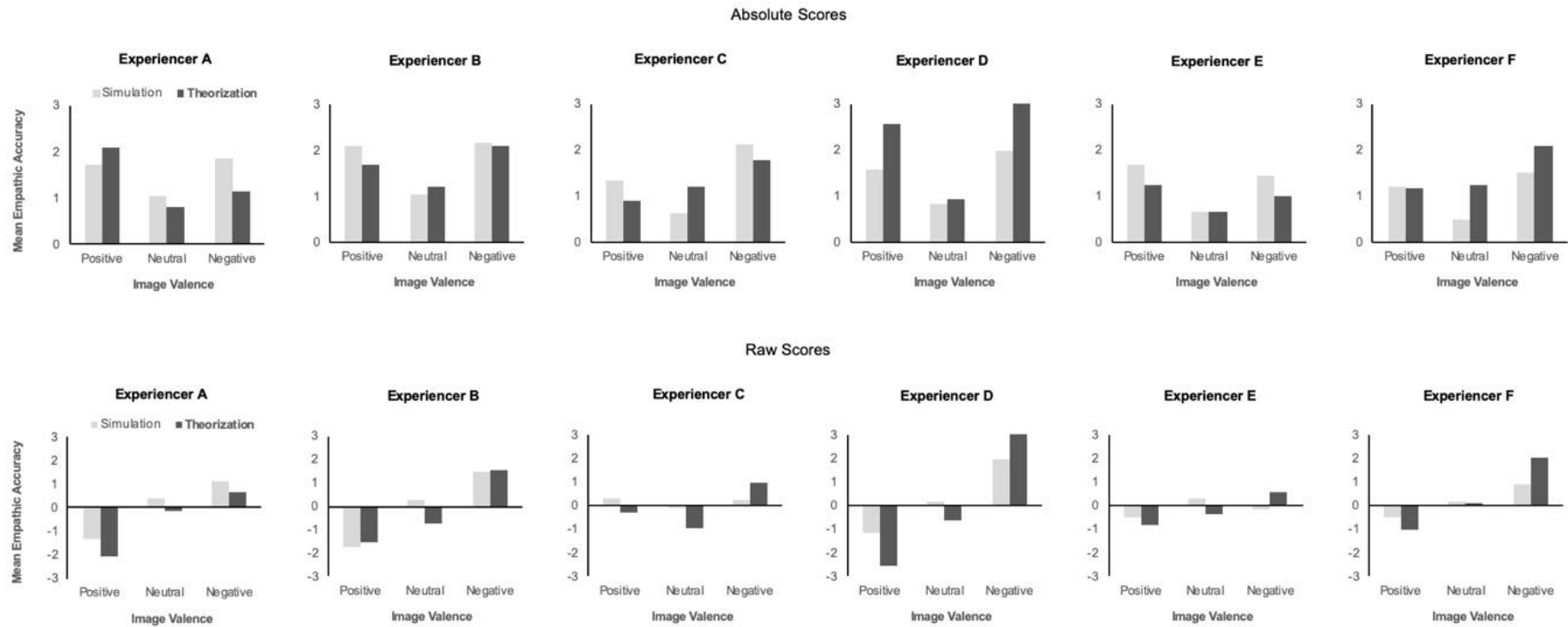
experiences.

Figure 9. Mean absolute (top) and raw (bottom) empathic accuracy scores (difference between experiencer scores and participant scores for experiencer) in the simulation (light grey) and theorisation (dark grey) groups for each of the six experiencers.

6.5 Discussion

6.5.1 Empathic accuracy varies as a function of empathic strategy

This study aimed to examine whether exposure to fiction is associated with mentalising or experience-sharing strategies for empathic accuracy. It attempted to replicate Zhou et al.'s (2017) findings that sharing in the experiences of targets (simulation) leads to higher empathic accuracy compared to mentalising based on facial cues (theorisation), and that participants tend to overvalue the theorisation strategy. In order to detect nuanced (positive and negative) variation in empathic accuracy, raw scores were calculated as well as absolute scores (the latter was in line with Zhou et al.). The absolute scores showed that the average magnitudes of error (differences between participants' ratings for the targets and the targets' actual ratings) were similar across the two conditions. However, the raw scores revealed that estimates in the theorisation condition tended to be over-negative whereas estimates in the simulation condition tended to be over-positive: when viewing experiencers' facial expressions, participants showed a negative bias, whereas when sharing in the targets' experiences by viewing the same set of pictures, participants showed a positive bias. Therefore, the present study did not replicate Zhou et al.'s finding that simulation represents the most effective strategy for accurately estimating the affective states of others (and did not support directional Hypothesis 5). Rather, each strategy was found to result in a similar level of error, although the direction of error differed.

Positive and negative biases for facial expressions have been well documented (for a review, see Kauschke et al., 2019). As negative bias (e.g., Baumeister, Bratlavsky, Finkenauer & Vohs, 2001; Rozin & Royzman, 2001) may serve an adaptive, predictive function, enabling perceivers to quickly identify a potential human threat (Hansen & Hansen, 1988), participants that viewed the facial expressions of targets may have been particularly predisposed to this bias. However, negative bias has also been documented with images (e.g., Pritsch

Telkemeyer, Mühlenbeck & Liebal, 2017) and yet it was not present in the simulation (picture-viewing) condition; in fact, it appeared to be reversed.

Two explanations may account for this finding: first, participants in the simulation condition may have moderated the initial negative bias via a process of reappraisal (Baumeister et al., 2001; Petro, Tong, Henley & Neta, 2018). In the theorisation condition participants were asked about (i) their emotional responses to the videos of the experiencers and (ii) their estimates of the experiencers' emotional ratings for the pictures. In contrast, in the simulation condition participants were asked about (i) their own emotional responses to the pictures and (ii) their estimate of the experiencers' emotional ratings for the pictures. Considering the same stimuli twice may facilitate a modification of the original estimate, and so reappraisal may represent a key process in the experience-sharing strategy. Further, as people tend to regulate their own negative emotions in order to minimise negative states (Baumeister et al., 2001), participant estimates may have been positively biased in the simulation condition where their own affective responses served as referents for the experiencers'.

Second, Zhou et al. (2017) pointed out that simulation may not represent the most effective strategy in all situations, rather it tends to be undervalued compared to theorisation, and this pattern is likely to be systematic. People's overconfidence in their ability to read facial expressions may lead to bias (the robust over-confidence effect; e.g., Pallier et al., 2002) whereas under-confidence in experience-sharing may minimise bias. While this could explain negative bias in the theorisation condition, it does not alone explain positive bias of equal magnitude in the simulation condition. The experimental paradigm is predicated on the assumption that the experiencers' perspectives on their initial emotional responses were accurate and not themselves subject to bias. If the experiencers' ratings reflected the initial negative bias, simulation estimates could be more accurate than the error magnitudes would

suggest. Future research may elucidate the effect of bias and the role of reappraisal in both experiencer and perceiver on empathic accuracy.

6.5.2 Empathic accuracy varies as a function of target and valence

Empathic accuracy varied as a function of target and valence and this differed between strategy conditions (supporting Hypotheses 6-8). Furthermore, almost all target $\times$ valence interactions were significant. Generally, estimated affect ratings for neutral pictures were the most accurate, and were the least accurate for negative pictures. Some experiencers' emotions were easier to detect than others, and this differed depending on whether the pictures were positive, negative or neutral. As identification with target did not relate to empathic accuracy, providing no evidence of an ingroup advantage (e.g., Adams et al., 2010; Matsumoto et al., 2009), it may be that particular experiencers expressed certain emotions more or less overtly than others. Research has indicated that the sex of the target can moderate the effect of valence on emotion labelling, for example Garrido and Prada (2017) found that participants correctly classified more "happy" expressions in static images of women and more "angry" expressions in men, though there was no interaction with "neutral" expressions. In the present study, participants showed generally higher levels of accuracy when rating the emotions of Experiencer E (male), although inspection of the target $\times$ valence interactions illustrates that accuracy with this experiencer was far higher for neutral images compared to both positive and negative images (Figure 9). It may be that individual differences in emotional expressiveness, the ability to accurately convey feelings nonverbally, impacted empathic accuracy in the present study. This tendency shows stability across contexts (Allport & Vernon, 1933) and is associated with other personality traits (positively with extraversion and negatively with neuroticism; Riggio & Riggio, 2002) as well as gender (women tend to be more nonverbally expressive than men; Hall, 1990).

Individual differences in the empathiser may also play a role in the ability to interpret positive, negative and neutral facial expressions across different experiencers. For example, individuals vary in "valence focus", the tendency to emphasise positive and negative information in verbal reports, which has been linked to an increased ability to process affective stimuli, and a heightened sensitivity towards negative facial expressions (Barrett & Niedenthal, 2004). Furthermore, the mood of the empathiser can bias recognition of incongruent emotions (e.g., a negative bias in participants primed with a sad mood and a positive bias in participants primed with a happy mood; Schmid & Schmid Mast, 2010). These findings reflect the importance of context when determining the most effective strategy for accurately interpreting the inner states of others. Future research may clarify the impact and interactions of individual differences in empathisers and experiencers on empathic accuracy for a range of emotions.

6.5.3 Fiction-exposure as a predictor of empathic accuracy

The correlation analysis revealed that fiction- and nonfiction-exposure were both negatively associated with absolute difference scores, indicating that the more people had read, the higher their empathic accuracy (supporting Hypothesis 9). Relationships were stronger for fiction-exposure in both conditions, in line with previous findings discussed in Chapter 3. Regression analyses revealed that fiction-exposure positively predicted variance in empathic accuracy even when nonfiction-exposure was controlled. This effect did not interact with condition, suggesting that while fiction-exposure is associated with empathic accuracy, it is not uniquely related to one empathic strategy over the other (and so Hypothesis 10 was not supported). It could be that both mentalising and experience-sharing strategies are enhanced, over time, through exposure to fiction, or that individuals with higher empathic skills tend to read more fiction.

6.5.4 Fiction-exposure as a predictor of empathic strategy choice

Zhou et al. (2017) found that participants over-valued theorisation, tending to self-select into this strategy condition despite simulation being the more effective approach to empathic accuracy. In the present study, accuracy in the theorisation and simulation conditions was comparable, yet theorisation also represented the most popular choice with the present sample. The strategy that had been previously used in the task (via the assigned condition) predicted strategy choice (in both conditions, participants tended to select the condition that they had not previously taken part in), but no other factors, including fiction-exposure, were found to significantly contribute to the prediction (and so Hypothesis 11 was not supported). This supports Zhou et al.'s suggestion that while the effectiveness of strategy may vary across contexts, there is a systematic tendency for people to overvalue their ability to read faces and undervalue insight gained from sharing in another's experience.

6.5.5 Limitations and recommendations for future research

As both absolute and raw difference values were employed to measure empathic accuracy, it is important to consider the potential error rate inflation. However, Bonferroni-adjustments for the two sets of dependent variables (raw and absolute scores) did not significantly alter the pattern of results (unless otherwise stated) and this approach provided a more nuanced view of the variance in empathic accuracy. By establishing the differences in empathic accuracy entailed by taking mentalising versus experience-sharing approaches, as well as the impact of contextual factors including the emotional content of the stimulus and individual differences between both experiencers and empathisers, the field will be better positioned to make recommendations for developing this integral social skill.

Identification with targets did not impact empathic accuracy in either condition. It may be that the perceived similarity to target question did not reliably access the extent to which participants identified with each experiencer. Indeed, the general lack of relationship between

identification with the target and the difference between participants' own ratings for the video clip or picture and their ratings for the experiencers suggests that this may be the case. Therefore, other factors may have led to the variation in empathic accuracy caused by different target experiencers, such as expressiveness or inhibition. Future research may develop this line of enquiry by examining effects across strategies. Participants were asked about which strategy they thought would be most useful after completing the task in their assigned condition, and so responses may have been biased by participants' experience of using the strategy assigned to them. Indeed, strategy used in the task was the only significant predictor of strategy selected after the task. Future research could examine strategy choice before the task is completed to obtain a purer measure.

Empathic accuracy was calculated in this study by computing the raw differences between participants' ratings for the target and the target's own ratings. The experimental paradigm, therefore, relies on the experiencers' abilities to accurately interpret their own affective responses. Employing dynamic facial reactions as stimuli represents a more ecologically valid method for assessing empathic accuracy compared to static photographs or schematic faces (e.g., Dobs, Bülthoff & Schulz, 2018), though the trade-off is that experiencers' self-reports could be subject to biases. Future research additionally employing validated behavioural tasks may help to establish systematic differences in empathic accuracy between mentalising and experience-sharing processes.

Finally, the tests concerning fiction-exposure yielded relationships rather than causal effects. Fiction-exposure appeared to predict empathic accuracy, though it is unclear whether it improves this skill or whether people with higher levels of this skill tend to read more fiction. Further research is required to identify the causal effects of fiction on the selection and value of each inferencing strategy.

6.5.6 Summary and moving forward

This study tested the respective value of mentalising versus experience-sharing strategies for accurately inferring others' affective states. It examined whether fiction-exposure predicts empathic accuracy, and the moderating role of inferencing strategy. Empathic accuracy was found to vary as a function of individual differences and situational aspects including valence, the identity of the target, and strategy used. It was positively predicted by fiction-exposure, although the effect of fiction-exposure was not impacted by mentalising or experience-sharing strategy.

Analysis of the raw scores revealed that mentalising and experience-sharing entailed similar levels of error, but that the direction of error diverged: mentalising based on facial cues resulted in negative bias, whereas experience-sharing—using one's own affective experience as proxy for the targets'—led to positive bias. It was suggested that reappraisal processes may play a role in accurately interpreting the affective states of oneself and others. Their impact on both mentalising and experience-sharing strategies for empathic accuracy could be fruitfully examined through future research.

This represents the first study to examine the relationship between fiction-exposure and mentalising and experience-sharing task performance. While the effects of condition diverged, condition did not moderate the positive impact of fiction-exposure on empathic accuracy. This indicates that fiction-exposure processes may support both mentalising and experience-sharing routes to empathic accuracy. However, the findings for fiction-exposure did not enable causal conclusions to be made, and so further research is required to examine causal effects of fiction on the empathy components. Furthermore, the empathic accuracy measure relied on the experiencers' abilities to accurately reflect on their own mental states and so it would be beneficial to examine fiction effects using validated behavioural measures of empathic accuracy.

The next and final study in this thesis aimed to examine causal effects of fiction on empathic accuracy using established empathic accuracy measures. It aimed to investigate one process through which fiction may enhance performance on empathic accuracy tests: narrative immersion. If stories are a simulation of the social world, becoming deeply immersed may mediate the effects of fiction on tests of empathy. As discussed in Chapter 3, previous research has indicated associations between immersion (or related constructs) and empathic accuracy. The final study in the present research project aimed to examine the effects of immersion in a story on empathic accuracy using an experimental approach.

Chapter 7: An Experimental Investigation of the Effect of Narrative Immersion on

Empathic Accuracy After Reading

In the previous two chapters, research was presented that established correlations between fiction-exposure and self-report (Chapter 5) and behavioural (Chapter 6) empathic ability measures. The present study extends this research by examining the causal effects of fiction-engagement, and specifically by testing the effect of one mechanism through which fiction may impact empathic accuracy: immersion. In line with the multidimensional approach to measurement taken throughout this research, this study incorporated multiple measures of immersion in order to tap into different aspects of the phenomenon. In an exploratory pilot study ($N = 61$), the experimental manipulation failed to vary levels of identification with characters, empathic feelings during reading and transportation into the storyworld. A re-design was undertaken which involved a materials pretest, followed by two manipulation pilots and a final experiment (total $N = 163$). The pretest and pilots established an immersive text and a cognitive load manipulation strategy for varying immersion levels. However, in the main experiment, the manipulation did not successfully vary immersion levels, and so a correlation analysis was undertaken which revealed positive associations between immersion measures, self-report and task-based empathic ability measures. An exploratory analysis examined the general prediction that reading would enhance empathic accuracy scores. This was found to be the case for one empathic accuracy measure (empathic accuracy for story characters' inner states), when controlling for demographic differences, trait empathy, fiction- and nonfiction-exposure, $p = .040$. This chapter presents the full set of five discrete experiments that contributed to the research agenda. It closes with a discussion of the methodological challenges involved in the study, and makes recommendations for future research aiming to establish the nature of fiction effects on empathic abilities, and the mechanisms through which such effects may occur.

7.1 Introduction

Empathic accuracy, the ability to accurately infer other people's mental states (Ickes, 1997), is essential to interpreting and responding appropriately to their thoughts, emotions and behaviours. Research has shown that the same cognitive and neural mechanisms are involved in the processing of both real and fictional events (e.g., Speer et al., 2009; Wallentin et al., 2011), suggesting that fictional narratives have validity as representations of the social world (Oatley, 2012; see also Gerrig, 1993; Zwaan, 2004). Through simulating the social experiences depicted in fiction, readers may develop their empathic inferencing abilities, leading to higher real-world empathic accuracy (e.g., Mar & Oatley, 2008; Oatley, 1999).

The meta-analyses described in Chapter 3 indicated that fiction-reading has an immediate, positive effect on empathic accuracy task performance (Dodell-Feder & Tamir, 2018) and that lifetime exposure to fiction positively correlates with empathic accuracy, as well as with self-report concern, experience-sharing and mentalising (Mumper & Gerrig, 2017). Furthermore, studies have shown that associations between fiction-reading and empathic abilities sustain when individual differences in general intelligence, age and English fluency, nonfiction exposure (e.g., Mar et al., 2006), openness to experience and the tendency to become absorbed in stories (Mar et al., 2009) are controlled.

7.1.1 Fiction and nonfiction

Social experiences are depicted in nonfiction as well as fiction narratives. Fictional narratives may uniquely induce empathy during reading because they enable the exploration of ideas and circumstances without the real-world obligations that may be implicated when reading nonfiction. In this way, fictional narratives may provide a "safe zone" in which to experience, infer and explore characters' inner states (Keen, 2007, p.4). While lifetime exposure to nonfiction is also positively associated with empathic accuracy, correlations tend

to be stronger for fiction and are observed when nonfiction-exposure is controlled (Mar et al., 2006; Mumper & Gerrig, 2017).

7.1.2 The "process" account

Stories may improve empathic abilities via the presentation of social information, or through the development of social processes, such as reasoning about mental states, perspective-taking and emotional engagement (Mar, 2018a). Most research evidence has supported the latter, "process", account (Mar, 2018a). For example, it has been suggested that literary devices such as striking aesthetic features and complex, unpredictable characters require readers to fill in "gaps" in the narrative, and entail an active, "writerly" approach to characters' inner experiences, which recruits and improves mentalising processes (De Mulder et al., 2017; Hakemulder, 2000; Kidd & Castano, 2013; Pino & Mazza, 2016; Zunshine, 2006). Such devices are primarily associated with "literary" fiction (fiction containing specific textual features such as foregrounding, e.g., Koopman, 2016; Kuzmičová et al., 2017, or that which has been previously classified as acclaimed or canonical, e.g., Kidd & Castano, 2013).

On the other hand, popular genre-fiction has also shown associations with empathic abilities (e.g., Fong et al., 2013) and moral reasoning (Black et al., 2018), and reading a suspenseful narrative has been linked to activity in areas of the brain associated with social cognition (Lehne et al., 2015). The first study in the current research project (reported in Chapter 5), showed positive associations between enjoyment of genre-fiction and empathic abilities (romance and drama genres showed small-medium positive correlations with self-report perspective-taking, empathic concern and fantasy, and experimental and comedy genres were positively associated with the same three IRI subscales and, additionally, with self-report altruism). Therefore, examination of the processes that mediate fiction effects,

which may transcend broad literary and popular classifications, would help to refine

knowledge of the relationships between fiction and empathic abilities.[19]

7.1.3 Narrative immersion

As outlined in Chapter 3, research in the field of persuasion has foregrounded the

impact of immersive process in narrative effects. Studies in this area have shown that fictional

stories can alter people's attitudes, beliefs and behaviours in predetermined directions (Slater,

2002), which can arise from the reader's analysis of explicit messages (analytic persuasion;

Chaiken, 1987; Petty & Cacioppo, 1986) or via the non-critical, experiential process of

becoming immersed, "lost" (Nell, 1988, p. 8) or "transported" (Gerrig, 1993) into a story

(narrative persuasion; Green & Brock, 2000; Green et al., 2004). Gerrig suggested that the

process of becoming transported into a storyworld may return readers to their original world

in some way altered, and Green and Brock (2000) posited transportation as the process

through which narrative persuasion occurs. They developed a measure of transportation

comprising eleven primary questions about quality of engagement, such as, "when I was

reading the narrative, I could easily picture the events in it taking place", and four additional

questions relating to story-specific content, for example, "while reading the narrative, I had a

vivid image of the boy" (Green & Brock, 2000, p. 704). Results showed that higher

transportation led to more story-consistent beliefs. For example, after reading a narrative

about the murder of a young girl by a psychiatric patient in an Ohio shopping mall, highly

transported readers were more likely to agree with story-consistent items such as there being a

high prevalence of violence in the USA and in shopping malls.

---

[19] In the study presented in Chapter 5, enjoyment of the factual/documentary genre was also positively associated with perspective-taking, empathic concern and altruism, although the associations were generally smaller than for the fiction genres. This was in line with previous studies showing that positive correlations between empathic abilities and nonfiction-exposure tend to be smaller than with fiction-exposure (Mumper & Gerrig, 2017). The narrative elements that mediate fiction effects may also transcend fiction versus nonfiction distinctions. Future research may examine the effects of varying the framing of narrative as fiction or nonfiction (for an example, see Green & Brock's, 2000, study of narrative persuasion).

Similarly, de Graaf, Hoeken, Sanders and Beentjes (2012) found that identification with a character during reading impacted attitudes after reading. Participants read a story told from the perspective of one of two characters with opposing points of view. Identification with the characters mediated the impact of perspective on attitudes after reading. If immersive processes such as transportation and identification with characters can facilitate the persuasive impact of implicit messages, such non-critical experiential processes may also facilitate empathy after reading.

7.1.4 Immersion and empathy

In Chapter 3, the conceptual overlap between immersion in stories and empathic processes was identified with reference to tools used to measure empathy. Gavaler and Johnson (2017, Experiment 1), for example, showed that adding mental state explanations to narratives increased transportation. As discussed in Chapter 2, the fantasy subscale of the IRI tends to be treated as a measure of trait immersion or transportability (e.g., Mar et al., 2006; see also Hall & Bracken, 2011) as all but one item refers to immersion in stories, mental imagery, and feelings towards fictional characters. Scores for this trait have been shown to positively correlate with performance on the IPT-15, a behavioural measure of empathic accuracy which, as described in Chapter 2, entails interpretation of filmed interactions between people ($r = .28$, $p < .05$; Mar et al., 2006). Fantasy also tends to correlate with the perspective-taking and empathic concern dimensions of the IRI ($r = .24$ and $.42$ respectively both ps $< .05$; Mar et al., 2006). The tendency to become absorbed in narratives appears to overlap with dimensions of empathy as diverse as the ability to accurately interpret facial expressions and the self-reported tendency to feel concern towards others. Evidence from neuroscience has also indicated an overlap between immersion and empathy, where studies have shown that areas of the brain associated with experience-sharing (the anterior and mid-

cingulate cortex; Walter, 2012) are activated during immersion in fiction (Hsu et al., 2014)

and continue to remain active after reading (Cadwell, 2015).

7.1.5 Causal effects of immersion

Behavioural research has further supported the link between immersion and empathy,

with some studies indicating that immersion in fiction may lead to increases in empathic

abilities after reading. Bal and Veltkamp (2013) found that emotional transportation

(measured using Busselle & Bilandzic's, 2009, scale) into a narrative increased self-reported

empathic concern a week after reading, whereas a lack of transportation appeared to reduce

empathic concern. However, this effect only held for participants that read fiction rather than

nonfiction (Study 1) and transportation was higher for fiction-readers (it may have been

impacted by the presentation of a single coherent narrative in the fiction condition, as opposed

to the selection of short, discrete texts presented in the nonfiction condition; Study 2).

Therefore, the effect on empathic concern could have been due to fiction-engagement

processes other than transportation.

Johnson and colleagues (Johnson, 2012; Johnson et al., 2013) provided evidence of a

positive effect of transportation on empathic feelings during reading and subsequent real-

world prosocial behaviour. Johnson (2012) found that transported readers experienced more

sympathetic emotions towards story characters (e.g., "compassionate"; Batson et al., 1987)

which, in turn, accounted for some increase in prosocial behaviour after reading. Johnson et

al. (2013) assigned participants to one of three groups and instructed them to (i) generate

imagery while reading, (ii) focus on verbal-semantic information (the meanings of particular

words within sentences), or (iii) read the text in the same way that they would usually read a

story for leisure. Participants who generated more imagery while reading the narrative became

more transported and were subsequently over three times more likely than participants in the

leisure reading group to behave in a prosocial way (by responding to a call for survey

participation by an ostensibly unfunded researcher for a nominal fee of $0.05). Here, the manipulation appeared to alter levels of transportation, although there was no manipulation check to assess how far instructions had been followed. Nevertheless, transportation fully mediated the effect of the imagery training on sympathetic feelings for the story characters (using the same measure from Batson et al., 1987, as the Johnson, 2012, study described above), and both transportation and sympathetic feelings for characters mediated effects on prosocial behaviour. This indicated that immersion in a narrative can promote prosocial behaviour both directly, and indirectly through increasing sympathy for story characters.

In contrast to these findings, in their study of the effects of reading literary fiction, Kidd and Castano (2013) observed no correlation between transportation and empathic accuracy task performance. Specifically, there was no relationship between transportation and cognitive mentalising (reasoning about thoughts and beliefs), although the measure, a false-belief task, suffered ceiling effects (Experiments 1 and 2). There was also no effect on interpreting emotions presented in pictures of faces (measured using the RMET, Experiments 1, 3, 4 and 5 and the DANVA-AF, Experiment 2). Transportation was not the focus of their study, however, and so it was not experimentally manipulated, and other effects of the reading manipulation may have attenuated any relationships. While immersion in fiction and empathy appear to be related, further research is required to establish if, and how far, immersion causally influences empathy after reading.

7.1.6 Narratives as stimuli

In Chapter 3, it was proposed that understanding the impact of immersion in narratives could have implications for understanding fiction effects, as well as for methods used to test empathic abilities. In narrative tasks (such as the SST and MASC described in Chapter 2) empathic accuracy is operationalised as the ability to correctly interpret the inner cognitive and affective states of the story's protagonists. However, texts are not equally transporting to

all readers (Green & Brock, 2000), and individual differences in identification with the characters, and the production of mental imagery and emotions during reading, may account for or attenuate variance in empathy task scores (see also Mar, 2018b). An investigation of immersion effects on empathic accuracy would not only contribute to knowledge of fiction effects, it could also inform the use of story stimuli in the field of empathy research.

7.1.7 Manipulating immersion

Immersion is a multifaceted process. It involves character identification, vicarious experiencing of characters' personalities and emotions, seeing events from characters' perspective and assuming their goals (Cohen 2001; Busselle & Bilandzic, 2008; Iguarta, 2010; Oatley, 1994, 1999; Slater & Rouner, 2002), attentional focus, mental imagery, and emotional engagement (Green & Brock, 2002). While there are overlaps between these aspects of narrative engagement, there are also important differences. For example, Green and Brock's (2000) definition of transportation includes an emotional aspect but does not require this to be related to characters, and identification with characters can extend beyond direct engagement with the narrative itself (de Graaf et al., 2012).

Immersion is multidimensional but its constituent constructs are rarely studied simultaneously. Research that has examined the effects of immersive processes on empathic abilities after reading (such as the studies described above) have focused on transportation, following Green & Brock's (2002) operationalisation of the construct. However, immersive processes may differentially impact fiction effects on components of empathy. It would be advantageous to take a multidimensional approach to both narrative immersion and empathic accuracy in order to establish if, and how far, immersion in fictional narratives impacts accuracy for mental state content.

Green and Brock (2000) successfully implemented a dual-task paradigm in order to manipulate transportation levels in participants during reading. The present study took a

similar approach to manipulating immersion levels in order test the hypothesis that high

immersion while reading fiction, compared to low immersion, would lead to higher explicit

empathic accuracy for cognitive and affective content, as well as a greater likelihood of

spontaneously inferring mental states (Hypotheses 12-15 articulated in Chapter 4). It aimed to

test the effect of this manipulation on different dimensions of immersion, and to identify

effects on different measures of empathic accuracy, in order to clarify the influence of

immersion in fiction effects on empathy.

7.2 Experiment 1 Method

7.2.1 Participants

Sixty-one social sciences students (84% female, aged 18-57, $M = 25.2$, $SD = 8.6$)

volunteered to participate in a study "investigating how people read stories" in return for

course credit. Participants were native English speakers (67%) or reported a minimum of 3

years' English fluency ($M = 9.05$, $SD = 4.37$).

7.2.2 Materials

7.2.2.1 Fictional story. *The End of Something* (Hemingway, 1988) is a short story (1424

words) that forms the stimulus text for the SST (Dodell-Feder et al., 2013; described in

Chapter 2). The narrative depicts two protagonists, Nick and Marjorie, out fishing together.

They talk, argue, and break up. Marjorie leaves and Nick's friend Bill appears. Nick and Bill

engage in a brief conversation about the relationship having ended, which indicates that Bill

had known about the breakup being forthcoming. The narrative contains descriptions of the

environmental setting, the characters' physical behaviours, and dialogue between them. Two

of the original character names were modernised (Marjorie was changed to Mandy and Bill

was changed to Will) and contextually specific terms were changed (e.g., "lumber schooners"

was changed to "lumber boats"). The Flesch Reading Ease Score (FRES; Flesch, 1948)

quantifies ease of reading comprehension on a scale of 0-100 (higher scores reflect easier

texts), and the score for the text was 93, indicating that it was very easy to comprehend. The Flesch-Kincaid Grade Level (FKGL) estimates the education level at which a text should typically be understood and a value of 2.7 indicated that the text could be read by students in UK Key Stage 2 (US Grade 3). No participants reported having read the story before.

7.2.2.2 Empathic accuracy.

7.2.2.2.1. Explicit and spontaneous mental state inferencing. The SST (Dodell-Feder et al., 2013) is a narrative-based test of explicit and spontaneous first- and second-order empathic accuracy (described in Chapter 2). Participants are asked to read *The End of Something* (above), and subsequently to answer questions about the characters and narrative and provide a summary of the story. The characters' inner thoughts and feelings are not made explicit in the story, and so high scores on the SST reflect the ability to synthesise contextual, verbal and physical information. Spontaneous mental state inferencing is measured as the presence or absence of a mental state inference in the story summary (first question). Then, eight questions probe explicit mental state inferencing, for example, 'Why is Nick afraid to look at "Mandy?", and five non-mental state (control) questions probe story comprehension. The data were scored using the simple rubric laid out by Dodell-Feder et al. Spontaneous empathic inferencing was scored as a dichotomous variable (present versus absent), and all other responses were scored on a scale of 0-2 for accuracy and complexity.

7.2.2.2.2 Cognitive and affective mentalising. The Yoni Test (Shamay-Tsoory & Aharon-Peretz, 2007) measures cognitive (24 trials) and affective (24 trials) first- and second-order mentalising and is sensitive to variation in healthy adults (see the description of the task in Chapter 2). The test features a central character, Yoni, depicted as a simple schematic face or "smilie", surrounded by four other images from a single category (e.g., cartoon animals, toys or photographs of people's facial expressions). Sentences appear at the top of the screen, and participants are asked to select the appropriate image to fill in the blanks in the sentences

(e.g., "Yoni loves BLANK"). In the first-order condition, participants must interpret Yoni's eye gaze and facial expression; in the second-order condition, participants must additionally interpret the gaze and expressions of surrounding characters, also depicted as schematic faces (e.g., "Yoni is thinking about the toy that BLANK loves). Sixteen control trials require interpretation of spatial relationships between Yoni and the surrounding objects (e.g., "Yoni is close to BLANK").

7.2.2.3 Narrative immersion. Narrative immersion was measured using three scales tapping into aspects of the phenomenon. The Transportation Scale (Green & Brock, 2000, 2002) comprises 11 general questions about vividness of cognitive attention to the story, imagery, feelings of suspense, emotional involvement and lack of awareness of real-world surroundings (e.g., "while I was reading the narrative, I could easily picture the events in it taking place"), plus questions about the characters and plot of the specific story. For the present study, five story-specific questions were included (e.g., "I had a vivid mental image of Nick"). Participants indicated agreement with statements from "not at all" to "very much" on a 7-point Likert-style scale. The internal consistency estimate was good with this sample (α = .86). Affective empathy (feelings during reading) was measured using the Empathy Index (Batson et al., 1987). Participants were asked to indicate how far six words, "softhearted", "compassionate", "moved", "tender", "warm" and "sympathetic" reflected how they felt while reading a story, in this case, the Hemingway text. Agreement was indicated on 5-point Likert scales from "not at all" to "extremely". Six irrelevant items (e.g. "intelligent") were included. Internal consistency was high (α = .89). The Identification with Character Scale (Iguarta, 2010) is a 14-item scale comprising empathic (cognitive and affective empathic reactions to characters, e.g., "I was worried about what was going to happen to the characters"), and merging (the experience of becoming a character and losing self-awareness, e.g., "I had the impression I was really experiencing the story of the characters") dimensions.

Agreement with statements was indicated on a 5-point Likert scale from "not at all" to "very much". Internal consistency was high ($\alpha = .94$).[20]

7.2.2.4 Trait empathic abilities. Three 7-item subscales from the IRI (Davis, 1980) multidimensional measure of trait empathy described in Chapter 3 were used: empathic concern, perspective-taking and fantasy (as previously mentioned, the personal distress scale reflects egoistic rather than altruistic responses to others in need, distinguishing it from the other subscales and so it was not used in this study). Empathic concern measures the trait tendency to feel concern and compassion for others ($\alpha = .74$), fantasy measures the tendency to become absorbed in narratives ($\alpha = .77$), and perspective-taking measures the tendency to reason about others' mental states ($\alpha = .68$). Deletion of one item on the perspective-taking scale, "if I'm sure I'm right about something, I don't waste much time listening to other people's arguments" would have increased alpha to .76. However, alpha tends to be lower than true reliability (e.g., Peterson & Kim, 2013), and as this was exploratory research using an established scale with a fairly small sample, the item was retained. Fourteen irrelevant items (e.g., "I have no difficulty in cooking a complicated meal") were included to mask the aims of the study.

7.2.2.5 Fiction-exposure. Acheson et al.'s (2008) version of the ART (see Chapter 3) was used ($\alpha = .94$). This version contains 65 fiction author names (e.g., "Alice Walker") and 65 plausible foils (composed of made-up names, e.g., "Beatrice Dobkin", or names of people who are not authors e.g., "John Landau"). Participants were required to select the names that

---

[20] A principle components analysis (PCA) was run on the narrative immersion measures to assess whether the three measures were distinct from one another. The KMO measure was .80 which indicated "meritorious" sampling adequacy (Kaiser, 1974) and Bartlett's test of sphericity was significant ($p < .001$) indicating that the data were factorisable. The analysis revealed a complex factor structure. Eight components accounted for 76.6% total variance and a forced three factor model was run which accounted for 57.2% variance, both with several cross-loadings. This indicates overlap across the constructs. Exploratory regression models were run as part of the procedure with transportation, affective empathy and identification computed as a single (additive) variable; however, the pattern of results was the same as those reported in the main regression analysis.

they knew to be fiction authors. In order to discourage guessing, participants were given the following instruction:

> Below is a list of names. Some of them are authors and some of them are not. Please select the ones that you know for sure are authors. Points are deducted for guessing, so you should only select those you are absolutely certain about. You can work quickly, clicking only those that you know. If you are unsure, just move on to the next name. Please answer truthfully, and do not try to look up the answers. It does not matter whether you know any or not—we are interested in finding out any names in particular that you have experience of.

7.2.2.6 Genre experience. In line with Study 1 (Chapter 5), participants were asked to indicate their tendency to read from the same set of seven genres (comedy, romance, crime/thriller, sci-fi/fantasy, drama/realistic, classical literature and nonfiction) on 5-point Likert scales anchored with "not at all" and "a great deal".

7.2.2.7 Demographic questions. Participants were asked to indicate their gender identity, age, English fluency and highest education qualification.

7.2.3 Procedure

Participants completed the demographic questions, trait empathic ability questions (IRI subscales), genre questions and fiction-exposure task, and were randomly assigned to one of two groups (immersion, distraction) to read the fictional short story. Instructions for reading varied between the groups: participants were asked to (i) read the story and try to imagine the characters and surroundings (immersion condition) or (ii) read the story while noting down words that would be difficult for a pupil in school year 5 (aged 9-10) to understand (distraction condition). This approach was based on that used in Green and Brock's (2000) Experiments 2-4, where some participants were instructed to identify segments of text that would be difficult for fourth-graders (UK year 5) to understand, which was designed to reduce transportation. After reading, participants completed the SST questions first as the stimulus forms part of the SST measure. These were followed by the narrative immersion scales

presented in a random order, after which the Yoni test was completed, as it was the only

measure that did not require memory of the text. Items were randomised within tasks, except

for the SST, wherein the order of questions is predetermined. All measures were administered

electronically, the empathic accuracy test via Inquisit and all other measures via Qualtrics.

The stimulus text was provided as a printout as immersion has been shown to be lower for

text presented on a digital screen (Mangen & Kuiken, 2014). Participants in the distraction

condition were additionally provided with a pen and paper for their notes.

7.2.4 Data analysis

Reliability analyses were conducted using Cronbach's alpha (coefficients are reported in

the Measures section). Pearson's r was used for all inter-scale associations, and the point-

biserial correlation coefficient (r_{pb}) was used for all correlations involving the dichotomous

spontaneous mentalising variable, with the exception of the gender and SST spontaneous

empathic accuracy correlation, which used Fisher's exact test (Fisher's exact test was reported

rather than Pearson Chi-Square, as one expected cell count was less than 5 at 4.1). Gender

differences and the effect of the manipulation on immersion scales were assessed using

independent t-tests. Following a null finding for the effect of the manipulation (reported

below), regression analyses were conducted to examine relationships between levels of

immersion and the empathic ability outcomes measures. Spontaneous mentalising was

predicted using binomial logistic regression, and multiple linear regression was used to

predict all other dependent variables. All statistical tests were two-tailed, and confidence

intervals were bias-corrected and accelerated using bootstrapping ($N = 1000$).

7.3 Results

7.3.1 Preliminary analysis

7.3.1.1 Group differences. Independent t-tests showed no significant differences in trait

levels of empathic concern, perspective-taking or fantasy between the two conditions, all ps >

.376 (the assumption of homogeneity of variances was met for all tests). Mean differences are presented in Table 9.

7.3.1.2 Gender differences. There was a statistically significant effect of gender on the SST explicit inferencing measure, with females ($n = 51$; $M = 10.24$, $SD = 2.27$) scoring higher than males ($n = 10$; $M = 7.1$, $SD = 3.73$), $t(59) = 3.56$, $p = .001$, 95% CI [1.37, 4.90]. The assumption of homogeneity of variances was met in all cases as indicated by non-significant results for Levene's test, $p < .05$. Males and females differed in genre enjoyment only in the case of the romance genre, which females preferred significantly more than males, $p = .015$. There were no gender effects on the SST spontaneous mentalising measure, $p = .075$, fiction-exposure, SST story comprehension, transportation, affective empathy or character identification, all ps > .23, or on any of the Yoni test dimensions, all ps > .14.

Table 9

Mean Scale Scores (and Standard Deviations) Averaged Across Items and Arranged by Condition.

	Empathic Concern	**Perspective-taking**	**Fantasy**	**Transp.**	**Affective Empathy**	**Identification**
Immersed	2.88 (0.41)	3.06 (0.58)	2.96 (0.59)	3.55 (1.05)	2.85 (0.84)	2.82 (0.94)
Distracted	2.98 (0.47)	3.08 (0.55)	3.03 (0.70)	3.62 (0.95)	2.97 (0.60)	3.00 (1.00)
Mean difference	$p = .376$ [-.33, .13]	$p = .885$ [-.31, .27]	$p = .651$ [-.40, .25]	$p = .760$ [-.59, .43]	$p = .608$ [-.55, .32]	$p = .477$ [-.68, .32]

Note. "Transp." = Transportation. Affective empathy and identification were averaged across 5-point scales; all other scales were 7-point scales. Immersed $n = 31$, Distracted $n = 30$, except for identification where Distracted $n = 29$. Bootstrapped ($N = 1000$) 95% bias-corrected and accelerated confidence intervals for the mean differences are presented in brackets.

7.3.1.3 Raw correlations. Inter-scale correlations are presented in Table 10. Significant correlations were generally positive, associations between the IRI subscales, the SST measures and Yoni test scales were small-medium, and associations between the narrative immersion scales were large in magnitude (Cohen, 1988). Trait fantasy was moderately

positively associated with transportation and character identification, empathic concern with affective empathy and character identification, and perspective-taking with all narrative immersion scales. Higher scores on the Yoni test physical trials were associated with a lower likelihood of making a spontaneous mental state inference on the SST. No other correlations reached significance. Although cognitive and affective scores on the Yoni test trials were related, neither was associated with the explicit or spontaneous SST questions. The exploration of genre was not relevant to the hypotheses and so these effects are not reported; however, self-reported tendencies to read from different genres were not associated with any empathic accuracy measures (all ps > .05).

   7.3.1.3.1 Assumption checks and transformations for correlation analysis. The assumption of homogeneity was violated for the point-biserial correlations for identification, explicit mentalising, and the Yoni test control condition, all ps < .041. It was met in all other cases as indicated by non-significant results for Levene's test. Applying a transformation to the variables of concern (squaring) removed outliers and heteroscedasticity for the character identification and explicit mentalising variables (Levene's test was non-significant at p = .130, and p = .079, respectively). The correlation between spontaneous and explicit mentalising remained significant $r_{pb}(59)$ = .31, p = .014, 95% CI [.06, .54], and between spontaneous mentalising and character identification remained non-significant, $r_{pb}(59)$ = -.006, p = .97, 95% CI [-.26, .26].

   Applying a rank transformation to the Yoni test physical trials data removed heteroscedasticity (Levene's test based on median, p = .090) and outliers, and the relationship between the Yoni physical trials and spontaneous mentalising remained non-significant, $r_{pb}(59)$ = -.20, p = .118, 95% CI [-.44, .05]. Box plots revealed outliers on the "yes" and "no" distributions of the fiction-exposure scale, the "yes" distribution of the perspective-taking

Table 10

Correlations between Trait Empathy, Fiction-exposure, Narrative Immersion and Post-test Empathic Accuracy Measures in Experiment 1.

	2	3	4	5	6	7	8	9	10	11	12	13
1. Fiction-exposure	.13 [-.07, .31]	.05 [-.19, .29]	-.07 [-.32, .16]	.24 [.01, .47]	.14 [-.14, .38]	.06 [-.15, .26]	.10 [-.09, .30]	.17 [-.05, .40]	.18 [-.04, .38]	.08 [-.14, .26]	.20 [-.006, .37]	.15 [-.11, .37]
2. Fantasy		.25* [.01, .48]	.45*** [.28, .61]	-.12 [-.39, .18]	.38** [.17, .58]	.18 [-.06, 39]	.42*** [.21, .60]	.07 [-.18, .31]	.15 [-.10, .40]	-.016 [-.21, .22]	.014 [-.21, .26]	-.009 [-.28, .25]
3. Perspective-taking			.32* [.10, .51]	.03 [-.23, .30]	.49*** [.23, .69]	.47*** [.29, .62]	.48*** [.25, .65]	.04 [-.27, .33]	.09 [-.20, .34]	-.10 [-.32, .14]	.08 [-.20, .33]	.17 [-.08, .43]
4. Empathic concern				-.10 [-.33, .12]	.24 [-.01, .50]	.38** [.20, .55]	.33** [.07, .54]	-.12 [-.39, .15]	-.05 [-.30, .21]	-.13 [-.34, .12]	.10 [-.17, .34]	-.01 [-.28, .26]
5. Story comprehension					-.08 [-.34, .21]	-.089 [-.34, .18]	-.06 [-.35, .23]	.14 [-.12, .39]	.14 [-.10, .37]	-.037 [-.28, .19]	.26* [.01, .48]	.22 [-.04, .45]
6. Transportation						.66*** [.50, .78]	.78*** [.65, .87]	.051 [-.23, .30]	.073 [-.20, .31]	.15 [-.08, .39]	.03 [-.22, .26]	.05 [-.20, .29]
7. Affective empathy							.71*** [.57, .83]	-.17 [-.37, .03]	-.11 [-.33, .10]	.05 [-.19, .30]	.09 [-.23, .35]	-.01 [-.25, .26]
8. Character identification								-.003 [-.27, .24]	.04 [-.21, .26]	.06 [-.18, .29]	.102 [-.18, .36]	.08 [-.21, .34]
9. Yoni cognitive									.80*** [.65, .89]	.40*** [.18, .59]	.06 [-.19, .41]	-.13 [-.34, .11]
10. Yoni affective										.48*** [.25, .67]	-.004 [-.26, .32]	-.16 [-.38, .09]
11. Yoni physical											-.104 [-.32, .20]	-.26* [-.43, -.06]
12. SST: Explicit												.35** [.13, .53]
13. SST: Spontaneous												

Note. *N* = 61, * *p* < .05, ** *p* < .01, *** *p* < .001; 95% bias-corrected and accelerated confidence intervals are presented in brackets.

scale, and the "no" distribution of the Yoni test control condition. These were not due to data errors and were retained in the analysis. The Shapiro-Wilk test for the Pearson product moment correlations indicated deviations from normality on the "yes" distribution of the perspective-taking subscale, story comprehension, explicit mentalising, the Yoni test cognitive, affective and control conditions, all ps < .048, and the "no" distribution for the Yoni test control condition, p < .001. Due to these violations, robust correlations were also conducted using Spearman's rho (r_s). The pattern of results remained the same and, additionally, fiction-exposure significantly correlated with story comprehension, $r_s(61) = .29$, $p = .023$, 95% CI [.05, .48], and cognitive mentalising, $r_s(61) = .27$, $p = .034$, 95% CI [.02, .50].

7.3.1.4 Manipulation check. As shown in Table 9, the immersed and distracted groups did not differ on any of the immersion scales, all ps > .477 (the assumption of homogeneity of variances was met for all tests). The effect of condition remained non-significant when fantasy was included as a covariate, all ps > .22. In sum, the manipulation did not have the desired effect of varying immersion levels across any of the immersion dimensions.

7.3.2 Primary analysis

Because the manipulation failed, the planned hypothesis tests were not conducted (effects of condition on Yoni test scores, SST explicit and SST spontaneous scores). Instead, regression analyses were used to examine how far the narrative immersion measures predicted empathic accuracy after reading.

7.3.2.1 Short Story Task. The narrative immersion measures (transportation, identification, and affective empathy), story comprehension and gender were used to predict explicit empathic inferencing on the SST.

7.3.2.1.1 Assumption checks. Correlations between the transportation, affective empathy and identification scales could indicate multicollinearity, but tolerance and VIF statistics were

acceptable. The highest correlation in the matrix was .58 (transportation with identification), the Durbin-Watson statistic was satisfactory at 2.28, indicating independence of observations, and collinearity diagnostics were satisfactory (lowest tolerance = .32 and highest VIF = 3.09). Visual inspection of the histogram and P-P plot indicated that the data were normally distributed, and partial regression plots showed linearity and homoscedasticity. There were no standardised residuals or studentised deleted residuals ± 3 standard deviations. There were no problematic leverage values, with all values below 0.2 except for two at 0.21 and 0.22. These results were considered unproblematic in combination with Cook's Distance values which were all < 1, suggesting no highly influential points (Cook & Weisberg, 1982).

7.3.2.1.2 Multiple linear regression. The model was statistically significant, $F(5, 55) = 4.32$, $p = .002$, $R^2 = .28$, adj. $R^2 = .22$, though gender, $B = -3.33$, 95% CI [-5.54, -1.06], $p = .004$, and story comprehension, $B = 0.35$, $p = .013$, 95% CI [.06, .61], were the only significant contributors. The pattern of results was the same with the inclusion of the trait empathy scales, fiction-exposure, age and education; again, only gender and story comprehension added significantly to the prediction, which accounted for 17% variance in explicit mentalising, $F(11, 49) = 2.12$, $p = .036$, $R^2 = .32$, adj. $R^2 = .17$.

7.3.2.1.3 Binary logistic regression. Narrative immersion, story comprehension and gender did not significantly predict the likelihood of making a spontaneous mental state inference when summarising the story for the SST, $\chi^2(5) = 9.22$, $p = .1$, although gender, $B = -1.76$, $p = .012$, 95% CI [-3.39, -.81], was individually significant with females 17% more likely to make a spontaneous inference than males. No outliers were found, and the model was an adequate fit, indicated by a non-significant Hosmer and Lemeshow test result, $p = .95$. The addition of the trait empathy scales, fiction-exposure, age and education did not substantially change the result and no individual variables significantly altered the odds of making a mental state inference, $\chi^2(11) = 10.75$, $p = .465$.

7.3.2.2 Yoni test cognitive.

7.3.2.2.1 Assumption checks. The Durbin-Watson statistic (2.12) indicated that the observations were independent. Collinearity diagnostics were satisfactory (lowest tolerance = .32, and highest VIF = 3.15), and the highest correlation was .57 (transportation with character identification). Visual inspection of the histogram and P-P plot indicated that the data were normally distributed and partial regression plots showed linearity and homoscedasticity. There were no standardised residuals or studentised deleted residuals ± 3 standard deviations. There were three leverage values just above the 0.2 "safe" threshold (0.21, 0.22 and 0.29), though these were well below the 0.5 "problematic" threshold, and Cook's Distance values were all below 1, suggesting no highly influential points.

7.3.2.2.2 Multiple linear regression. Narrative immersion, story comprehension and performance on the affective and control Yoni test trials significantly predicted Yoni cognitive scores, $F(6, 54) = 16.84$, $p < .001$, $R^2 = .65$, adj. $R^2 = .61$, though performance on the affective trials was the only significant contributor, $B = .85$, $p < .001$, 95% CI [.64, 1.06].

7.3.2.3 Yoni test affective. A second regression model substituted cognitive for affective mentalising.

7.3.2.3.1 Assumption checks. The Durbin-Watson statistic (2.17) indicated no issue with autocorrelation and collinearity diagnostics were satisfactory (lowest tolerance = .32 and highest VIF = 3.13). The histogram and P-P plot indicated that the data were normally distributed and partial regression plots showed linearity and homoscedasticity. There were no standardised residuals or studentised deleted residuals ± 3 standard deviations. There were three leverage values just above the 0.2 "safe" threshold (0.21, 0.23 and 0.25), though these were well below 0.5, and Cook's Distance values were all below 1, suggesting no highly influential points.

7.3.2.3.2 Multiple linear regression. The model significantly predicted affective mentalising scores on the Yoni test, $F(6, 54) = 18.69$, $p < .001$, $R^2 = .68$, adj. $R^2 = .64$, though only the Yoni test cognitive, $B = .63$, $p < .001$, 95% CI [.48, .80], and control trials, $B = .12$, $p = .024$, 95% CI [.012, .23], were significant contributors.

7.3.2.3.3 Inclusion of covariates. Including gender, age, education, trait empathy levels and fiction-exposure in the models did not significantly alter the pattern of results: only cognitive mentalising added significantly to the model for affective mentalising $F(13, 47) = 9.22$, $p < .001$, $R^2 = .71$, adj. $R^2 = .64$, and only affective mentalising significantly contributed to the prediction for cognitive mentalising $F(13, 47) = 7.96$, $p < .001$, $R^2 = .69$, adj. $R^2 = .601$. All assumptions were met for these models (the Durbin-Watson statistic was 1.98 for the model predicting cognitive mentalising, and 2.05 for the model predicting affective mentalising. No correlations were greater than .58 and the highest VIF value was 3.41).

7.3.2.4 Yoni test order. To assess whether scores differed for first- and second-order trials, a two-way repeated measures ANOVA was conducted to assess the effects of order (first, second) and trial type (cognitive, affective, control) on Yoni test scores. Results showed that the effect of trial type was non-significant, $F < 1$, though there was a significant main effect of order, $F(2, 60) = 27.53$, $p < .001$, $\eta_{p}^2 = .32$, and a significant interaction effect, $F(2, 59) = 7.86$, $p = .001$, $\eta_{p}^2 = .21$. Follow-up regression models were conducted to assess whether narrative immersion predicted first and second-order cognitive and affective Yoni test scores. All models (which included story comprehension) were non-significant, all ps > .10.

7.4 Discussion

The aims of the pilot experiment (Experiment 1 in this chapter) were to assess whether immersion in a narrative impacted explicit and spontaneous mental state inferencing (using the SST measure), and mentalising about cognitive and affective content (using the Yoni test).

The experimental manipulation failed to impact levels of immersion in the Hemingway story, the SST stimulus text. Therefore, regression models were used to assess how far immersion levels across three immersion scales predicted empathic accuracy. These findings did not support the prediction that higher immersion would relate to higher empathic accuracy.

7.4.1 Limitations

The results revealed strong correlations between self-reported empathic ability traits measured using the IRI and narrative immersion: transportation correlated with perspective-taking, and character identification and affective empathy during reading with both perspective-taking and empathic concern. This suggests that similar processes may underpin relationships with real-world and fictional others; thus, it remains possible that immersion impacts empathic accuracy, but that the present study failed to detect the effect. One reason for this could be the capacity of the stimulus text to evoke immersion. Ernest Hemingway, the author of the stimulus story, has an economical, non-decorative style, which may have accounted for the observed generally low immersion levels. Immersion scores were low across both groups (average total scale scores in both groups were below 50 for transportation, below 29 for identification with character and below 12 for affective empathy, compared to maximum possible scores of 96, 56 and 24 respectively. The transportation scores reported in Green and Brock's, 2000, experiments ranged from 31-99 of a possible 105). Additionally, the instruction designed to lower immersion levels may have encouraged perspective-taking, as participants were effectively tasked with considering the text from the perspective of a 9-10 year-old child, which could have attenuated the group difference. Finally, due to the lack of a no reading control group in the present experiment, it was not possible to detect any effects of reading a story on empathic accuracy. The following experiments were designed to address these limitations.

7.4.2 Moving forward

The next experiment aimed to retest the hypothesis that high immersion in a fictional narrative would lead to greater empathic accuracy compared to low immersion, while addressing the central limitations of Experiment 1: low immersion in the narrative across both experimental groups and potentially problematic instruction in the distraction group, and the lack of a no reading control group. The exploration of genre was superfluous to the main aim of the study and so in order to focus on the effects of immersion, questions about genre were not included in the next iteration of the study. Ahead of the main experiment, a pretest established the most immersive of three texts, and two pilot studies tested experimental procedures designed to manipulate immersion.

7.5 Experiment 2: Text Selection Pretest

A pretest was conducted to identify an immersive text by establishing effects of different stories on the dimensions of immersion. Two new texts were added for comparison with the Hemingway story used in the pilot: a passage from a popular fiction novel and a short story written for the study. The study aimed to compare the effects of the three texts on immersion, while controlling for trait fantasy as an index of the general tendency to become immersed in stories.

7.5.1 Method

7.5.1.1 Participants. Fourteen participants (10 females, 4 males) completed the study online via a participant recruitment website (Call for Participants) in return for a £10 online shopping voucher. The study was presented using Qualtrics. One participant was excluded for incomplete data (indicating that they withdrew from the study) and another two were excluded for reporting having read one or more of the texts prior to participating, resulting in $N = 11$ (8 females) aged 25-58 ($M = 33.45$, $SD = 8.81$). All participants spoke English as their main language.

7.5.1.2 Materials.

7.5.1.2.1 Stimulus texts. Three fictional texts were chosen for the study: a literary text, a passage from a popular fiction novel, and a short story written by the researcher. The literary text was Ernest Hemingway's (1988) *The End of Something*, which forms the stimulus text for the SST (Dodell-Feder et al., 2013) and was used in the pilot study. The popular fiction text was a section of *The Unlikely Pilgrimage of Harold Fry*, a bestseller by Rachel Joyce (2012), in which the central character, a retired man called Harold, decides to journey across England on foot in order to visit an old acquaintance whom he learns is dying. In the selected segment, (pp. 22-29, 2006 words), Harold spontaneously decides to begin his journey. He visits a local petrol station where he purchases food and interacts with a sales assistant. Like *The End of Something,* the focalisation is third-person ("Harold did…", rather than "I did…"). This section was selected as it is a similar length to *The End of Something,* appears early in the novel and so does not require additional, contextual knowledge of the story, and because it contains descriptions of the environment and of the protagonist's inner thoughts, which may support immersion. As discussed in Chapter 3, the inclusion of mental state references has been associated with higher levels of immersion in narratives (Gavaler & Johnson 2017, Experiment 1) and the lack of mental state information in the Hemingway text may have contributed to the low immersion levels observed in the pilot study. The final text, *The Wallet* (1788 words), is a short story written for the study (Appendix F). The narrative was written from the first-person perspective as this focalisation has been shown to enhance identification with character (de Graaf et al., 2012). The narrative follows a protagonist's journey to work during a cold morning. On the way, the protagonist finds a wallet and makes efforts to return it to its owner despite their discomfort in the cold weather and the consequence that they will be late for an important work meeting. The outcome of the story is positive: the owner of the wallet shows gratitude and the protagonist's employer hints at promotion. In other words, the

protagonist performs an altruistic act, for which they are rewarded. In order to elicit immersion, the narrative included descriptive passages depicting the surrounding environment, and detail about the protagonist's inner states (from a first-person perspective). The protagonist remains unnamed throughout the text, and descriptions were constructed so as not to signify a specific gender or cultural identity, in order to encourage identification.

7.5.1.2.2 Narrative immersion. The same three scales were used to measure immersion as in the pilot study: the Transportation Scale (Green & Brock, 2002; $\alpha = .84$), the Empathy Index (Batson et al., 1987; $\alpha = .96$), and the Identification with Character Scale (Iguarta, 2010; $\alpha = .94$).

7.5.1.2.3 Fantasy. The fantasy subscale of the IRI (Davis, 1980), used in the pilot, was included as an index of the trait tendency to become immersed in stories (e.g., Mar et al., 2006; Hall & Bracken, 2011; $\alpha = .83$).

7.5.1.3 Procedure. After entering their demographic details, participants were asked to complete the fantasy subscale of the IRI. They were then asked to read the three fictional passages, which were presented in a random order. After reading each story, participants were asked to complete the three narrative immersion measures, which were internally randomised and presented in a random order.

7.5.1.4 Data analysis. The effects of story on the three immersion variables were examined using three repeated measures analyses of covariance (ANCOVAs), partialling out variance caused by the trait tendency to become immersed in stories (fantasy). Descriptive statistics using scores averaged across items are presented to facilitate interpretation of the scales across the three reading groups (identification and affective empathy were averaged across 5-point scales, whereas transportation was averaged across a 7-point scale).

7.5.2 Results

7.5.2.1 The effect of text on immersion. Means averaged across each scale are presented in Table 11. Due to anticipated low power, pairwise comparisons were examined for non-significant main effects (Bonferroni adjusted values are reported, and the same pattern of results emerged with no adjustment).

7.5.2.1.1 Identification with character. Controlling for trait fantasy, the overall effect of stories on the identification with character scale was non-significant, $F(2, 18) = 2.04$, $p = .16$, $\eta p2 = .19$ (the assumption of sphericity was met, indicated by Mauchly's test statistic, $p = .062$). However, pairwise comparisons revealed a significant difference between the Hemingway story compared to the other two stories, such that total character identification scores for the Hemingway characters were lower the compared to characters in *The Wallet*, $p = .002$, and the Joyce characters, $p = .047$. There was no difference between identification scores for characters in *The Wallet* versus Joyce characters, $p = 1$.

Table 11

Overall Means (and Standard Deviations) of Scale Scores Averaged Across Items for Immersion Measures Presented by Story.

Story	Identification	Affective Empathy	Transportation
Hemingway	2.22 (0.88)	2.11 (0.28)	3.39 (0.24)
Joyce	3.01 (0.83)	3.20 (1.19)	4.52 (0.75)
The Wallet	3.08 (0.94)	2.61 (1.00)	3.95 (0.99)

Note. Affective empathy and character identification were measured on 5-point Likert scales, and transportation was measured on a 7-point Likert-style scale. Values greater than 3 are on the positive side of the scale for affective empathy and character identification, but values greater than 4 are on the positive side of the transportation scale.

7.5.2.1.2 Affective empathy. Controlling for fantasy, the overall effect of story was non-significant, $F(1.26, 11.34) = 1.04$, $p = .35$, $\eta p2 = .10$ (the assumption of sphericity was not met, indicated by Mauchly's test statistic, $p = .029$, and so Greenhouse Geisser's adjusted F-ratio is reported). Pairwise comparisons revealed a significant difference between the

Hemingway and Joyce stories, with participants reporting lower levels of affective empathy for the Hemingway text, $p = .008$. No other comparisons were significant.

7.5.2.1.3 Transportation. Controlling for fantasy, the effect of story on transportation was non-significant, $F(1.27, 11.40) = 1.94$, $p = .19$, $\eta p_2 = .18$ (the assumption of sphericity was not met, indicated by Mauchly's test statistic, $p = .032$, and so Greenhouse Geisser's adjusted values are reported). Again, pairwise comparisons revealed significant differences between stories. Transportation scores for the Hemingway story were significantly lower than for the Joyce story, $p < .001$, and no other comparisons were significant.

7.5.3 Moving forward

The popular fiction passage from the Joyce story was found to have the most impact on the immersion scales overall and so it was selected to form the stimulus text for the full experiment. The next experiment was designed to pilot an adaptation of the original manipulation, which failed in the first experiment, with the newly selected immersive text.

7.6 Experiment 3: Distraction Task Manipulation Pilot

The manipulation pilot was designed to test the between-participants effect of an adaptation of the dual task paradigm. This was designed to reduce immersion levels in one group without inducing participants to take the perspective of another individual or group (as was a potential confound in Experiment 1).

7.6.1 Method

7.6.1.1 Participants. Thirty-eight participants were recruited via the same online participant recruitment website as the text pretest (Call for Participants) and were given the option of entry into a prize draw to win a £10 online shopping voucher in return for participating. Three participants were excluded for reporting a main language other than English, four were excluded for reporting having read the story prior to participating in the study, and one for failing to respond correctly to a basic story comprehension question

(requiring correct selection of the central characters' names). This resulted in a final sample of 30 participants (23 females) aged 21-66 ($M = 33.33$, $SD = 12.70$).

7.6.1.2. Materials. The stimulus text was the same passage from *The Unlikely Pilgrimage of Harold Fry* (Joyce, 2012) used in the text pretest. The same three immersion scales were used as in the pilot study and text pretest (Transportation Scale, $\alpha = .88$; Empathy Index, $\alpha = .87$; Identification with Character Scale, $\alpha = .92$). Again, trait fantasy was measured using the IRI fantasy subscale ($\alpha = .78$). Attention check questions were included with the immersion scales, which required participants to select a specific option (e.g., "if you are paying attention, please select scale point 3"). Participants were excluded from the analysis if inspection of the dataset revealed incorrect answers to any of these checks. Following the story, participants were asked two basic multiple-choice comprehension questions: "where did Harold buy a cheeseburger from?" (correct answer: a petrol station) and "who was Harold trying to contact?" (correct answer: Queenie). Skip logic was set up in Qualtrics such that participants were automatically taken to the debrief if they failed to answer either of these questions correctly and their data were flagged for exclusion.

7.6.1.3 Procedure. Participants first completed the demographic questions and fantasy scale. They were randomly assigned to one of two groups (immersion, distraction task) to read the fictional passage. Participants in the immersion group received the following instruction: "please take your time to read it all and allow yourself to become really absorbed in the story. Use your imagination to visualise the environment and the characters as much as you can. Try to put yourself in the position of the main character and imagine things from their point of view". Participants in the distraction task group received an alternative instruction: "please take your time to read it all and look out for examples of proper nouns. A proper noun is a name used for a person, place or organisation, spelled with a capital letter (e.g., Tim, America or Starbucks). After reading the story, you will be asked about these".

7.6.1.4 Data analysis. In line with the pilot and text pretest, immersion and fantasy scale scores were averaged across items to support interpretation of scores relative to scale points. *T*-tests were conducted to establish the effect of the manipulation on immersion levels and ANCOVAs were used when fantasy was included as a covariate. Confidence intervals were bias-corrected and accelerated using bootstrapping ($N = 1000$).

7.6.2 Results

7.6.2.1 Effect of manipulation. Average mean immersion levels for each group are presented in Figure 10. Immersion levels were low compared to the pretest levels. There were no significant effects of condition on transportation ($M_{immersion} = 4.27$, $SD = 0.91$; $M_{distraction} = 4.20$, $SD = 1.19$), identification with character ($M_{immersion} = 2.73$, $SD = 0.70$; $M_{distraction} = 2.56$, $SD = 0.80$), or affective empathy ($M_{immersion} = 3.02$, $SD = 0.65$; $M_{distraction} = 3.03$, $SD = 0.93$), all $Fs < 1$ (the assumption of homogeneity of variances was met for all tests indicated by non-significant Levene's test statistics, all $ps > .05$). Fantasy did not differ between conditions, $t(28) = .034$, $p = .97$, 95% CI [-3.74, 3.86], and inclusion of the fantasy covariate did not alter the pattern of results (for effects of condition, all $Fs < 1$).

7.6.3 Moving forward

The new distraction task manipulation failed to vary immersion in the stimulus text established in the pretest. Because participants either received an instruction designed to induce immersion, or a distraction task instruction designed to reduce it (effectively two different manipulations) and there was no control condition, it is difficult to assess which had the most impact on the way that participants engaged with the text. Nevertheless, there was a lack of difference between the two groups on all three immersion scales, and so an alternative approach was required to manipulate levels of immersion. This was the focus of the second manipulation pilot.
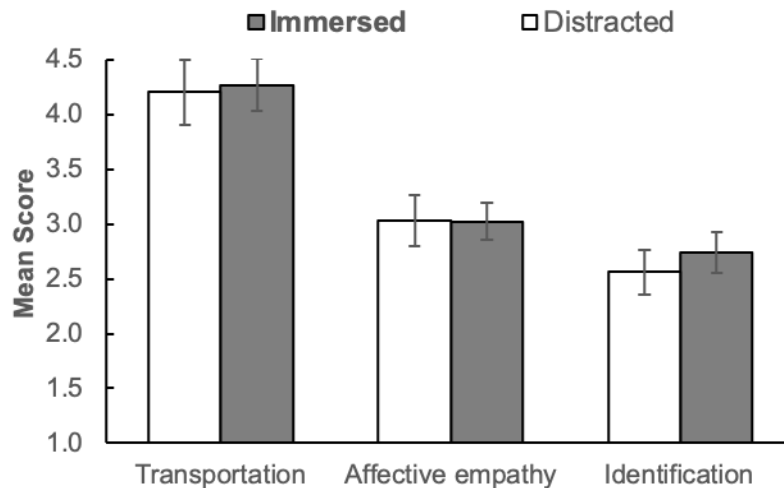
Figure 10. Raw transportation, affective empathy and identification scores presented for immersion and distraction conditions (Experiment 3). Affective empathy and character identification were measured on 5-point Likert scales, and transportation was measured on a 7-point Likert-style scale. Error bars are standard errors of the mean.

7.7 Experiment 4: Cognitive Load Manipulation Pilot

In their examination of the effects of transportation on narrative persuasion, Green and Brock (2000) used a dual task approach designed to reduce immersion levels in one group. A similar approach was used in the pilot (Experiment 1) and first immersion manipulation pilot (Experiment 3), where it failed to manipulate immersion levels. This necessitated an alternative strategy for varying immersion. Like the previous distraction task manipulation, a cognitive load task was also drawn from the narrative persuasion literature where it has been employed to distinguish between automatic and elaborative processing of text (e.g., Rocklage, Rucker & Nordgren, 2018; Tormala & Petty, 2004; Zemborain & Johar, 2007; see also Gilbert, 1991; Gilbert & Hixon, 1991). Green and Brock (2000) proposed that transportation is "qualitatively different" (p. 702) from the cognitive elaboration associated with dual process accounts of narrative persuasion (which contrast a central, elaborative route with a heuristic, peripheral route; e.g., Petty & Cacioppo, 1986; see also Lord, Lee & Sauer, 1995; MacKenzie, Lutz & Belch, 1986). However, based on Green and Brock's (2000) finding that completing an additional task while reading effectively reduced immersion, it was predicted

that the cognitive load manipulation would reduce levels of immersion, and hypothesised that this, in turn, would have an effect on post-reading empathic accuracy. The current experiment aimed to test the effect of the cognitive load manipulation on immersion levels.

7.7.1 Method

7.7.1.1 Participants. Fifty participants were recruited via the same participant recruitment website used in the previous experiments (Call for Participants). Links to the site were posted on university social media forums (Facebook and Twitter), and participants were given the option to be entered into a prize draw to win a £10 online shopping voucher. Fourteen participants were excluded either for failing attention checks ($n = 10$) or ballot box stuffing (completing the study multiple times; $n = 4$), and seven participants were excluded for having read the stimulus story prior to participating in the study. This resulted in total $N = 29$ (27 females) aged 18-66 ($M = 36.76$, $SD = 12.57$). Due to issues with attempts at ballot box stuffing, study completion durations were also inspected. All participants in the final sample completed the study within 3 standard deviations of the mean.

7.7.1.2 Materials. Materials were the same as those used in the previous immersion manipulation pilot (Experiment 3) and were presented in Qualtrics online.

7.7.1.3 Procedure. Participants completed demographic questions and the fantasy subscale of the IRI (internally randomised). They were assigned to either a high or low cognitive load condition and asked to remember an ostensibly random number: a standard cognitive load manipulation (Rocklage et al., 2018; Tormala & Petty, 2004; see also Gilbert & Hixon, 1991). Participants in the high cognitive load condition were asked to keep in mind an eight-digit number (31875649), and participants in the low cognitive load condition were asked to keep in mind a two-digit number (27), while reading the stimulus text. Instructions were the same for both groups of participants:

On the next page, you will be presented with a number. It will appear for 20 seconds before

the page moves on. Your task is to remember that number, as you will be asked to recall it

later. It's really important for the study that you don't write the number down or copy and

paste it, just try to keep it in mind.

After reading the story, all participants were asked to enter the number that they had been

required to keep in mind, and then to complete the three narrative immersion scales (internally

randomised and presented in a random order).

7.7.1.4 Data analysis. Scale scores were averaged across items (across a 7-point scale

for transportation and 5-point scales for identification and affective empathy). Scale reliability

was computed using Cronbach's alpha. Group differences were examined using *t*-tests and

ANCOVAs. Confidence intervals were bias-corrected and accelerated using bootstrapping (*N*

= 1000).

7.7.2 Results

7.7.2.1 Scale reliability. Cronbach's alpha values for this sample were .77 for the

fantasy scale, .95 for the identification with character scale, .93 for affective empathy and .76

for transportation.

7.7.2.2 Fantasy. The high cognitive load group showed higher levels of trait fantasy

(averaged across the scale, $M = 3.66$, $SD = 0.42$) compared to the low cognitive load group

($M = 3.11$, $SD = 0.87$), $t(26.92) = -2.26$, $p = .032$, 95% CI [-7.27, -0.35]. Consequently, the

analysis of the manipulation effect was conducted both with and without the fantasy

covariate.

7.7.2.3 Effect of manipulation. Average mean immersion levels for each load group are

presented in Figure 11. Without the inclusion of the covariate fantasy, *t*-tests (assuming equal

variances due to non-significant Levene's test results, $p > .05$) revealed that there were no

significant differences between the high and low cognitive load conditions on any of the three

immersion scales, all ps > .20. ANCOVAs including the fantasy covariate revealed significant

effects of condition across all scales (the assumption of homogeneity of variances was met in all cases indicated by non-significant Levene's test statistics, $ps > .05$). Identification with character was higher in the low cognitive load condition ($M = 2.91$, $SD = 1.04$) compared to the high load condition ($M = 2.81$, $SD = 0.83$), $F(1, 26) = 4.40$, $p = .046$, $\eta p_2 = .15$; affective empathy was also higher in the low load ($M = 3.26$, $SD = 1.05$) compared to the high load condition ($M = 2.80$, $SD = 0.95$), $F(1, 26) = 7.21$, $p = .012$, $\eta p_2 = .22$; and transportation was also higher in the low load ($M = 3.95$, $SD = 0.92$) compared to the high load condition ($M = 3.57$, $SD = 0.73$), $F(1, 26) = 8.95$, $p = .006$, $\eta p_2 = .26$.
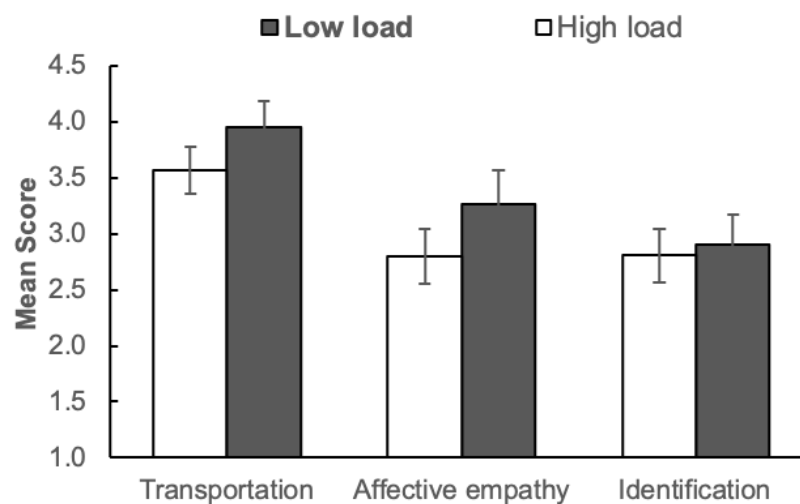
Figure 11. Raw transportation, affective empathy and identification scores presented for high and low cognitive load conditions (Experiment 4). Affective empathy and character identification were measured on 5-point Likert scales, and transportation was measured on a 7-point Likert-style scale. Error bars are standard errors of the mean.

7.7.3 Moving Forward

This experiment examined the effects of a cognitive load manipulation on immersion in the selected immersive text. The results showed that the low cognitive load condition (retaining a two-digit number) led to higher levels of immersion across all three measures compared to the high cognitive load cognition (retaining an eight-digit number) when variance caused by trait fantasy, which was not matched across the groups, was partialled out

(examination of the fantasy data revealed a greater spread in the low cognitive load group, where total scores ranged from 11-31 [$M = 21.79$, $SD = 6.12$], compared to the high cognitive load group where total scores ranged from 21-31 [$M = 25.60$, $SD = 2.95$]). Additionally, the high and low cognitive load group sizes were unequal ($n = 10$ and $n = 19$, respectively), and the sample size was small. Taking these factors into account, the cognitive load manipulation appeared an appropriate strategy for varying immersion during reading with a larger sample in an investigation of the effects of narrative immersion on empathic accuracy post-reading.

7.8 Experiment 5: Effects of Narrative Immersion on Empathic Accuracy

Manipulating cognitive load was shown to successfully vary immersion across all three scales when trait fantasy was controlled. The effect was in the expected direction: participants in the high cognitive load condition tended to report lower immersion levels compared to participants in the low cognitive load condition. Therefore, this approach was used to vary immersion levels with the selected immersive text in the main experiment, which aimed to examine the effects of immersion on empathic accuracy measures.

7.8.1 Method

7.8.1.1 Participants. A power analysis was conducted in G* Power (Faul et al., 2007). The smallest effect of cognitive load on identification with character, established in the second manipulation pilot ($\eta p_2 = .15$), was used in the computation. This indicated that a total sample of 49 (24-25 participants per group) would be required to detect the effect of the manipulation on immersion at 80% power, when controlling for trait fantasy. The small effect size of fiction on empathic accuracy (Dodell-Feder et al., 2018) and of immersion in fiction on empathic concern (Bal & Veltkamp, 2013) indicated that a much larger sample ($N > 193$) would be required to achieve 80% power to detect effects of reading and of immersion; however, this sample size was unattainable due to resource constraints. Data were gathered from 100 participants who were all undergraduate and postgraduate students of psychology at

Kingston University London. Participants received course credit for volunteering their time and were also offered the opportunity to enter a prize draw to additionally win a £10 online shopping voucher. It was a requirement of the study that participants spoke English as their main language. Four participants reported other main languages and one participant reported four main languages including English, and these participants were excluded from the dataset. Two participants were excluded for failing one or more attention checks (there were two forms of attention check question: simple questions about the story content and Likert-style questions explicitly requiring participants to select a particular response to show that they are attending to the questions, as described in the pilots). This resulted in a total sample size of 93 (75 females, aged 18-54, $M = 24.75$, $SD = 7.62$). Most participants had completed A-levels or equivalent (54%), 37% had completed a graduate degree, and 9% a postgraduate degree. Participants were assigned to either the high cognitive load ($n = 28$), low cognitive load ($n = 35$) or no reading ($n = 30$) conditions.

7.8.1.2 Materials.

7.8.1.2.1 Fictional story. The passage from *The Unlikely Pilgrimage of Harold Fry* (Joyce, 2012) established in the pretest (Experiment 2) and used in the manipulation pilot studies (Experiments 3 and 4), formed the stimulus text. No participants reported having read or heard of the story prior to participating.

7.8.1.2.2 Narrative immersion. The Identification with Character Scale ($\alpha = .94$), Empathy Index ($\alpha = .90$) and Transportation Scale ($\alpha = .80$) were used, as used in the pretest and pilots.

7.8.1.2.3 Empathic accuracy. The Yoni test and the SST were used to measure empathic accuracy, as described in the pilot study. The Yoni test measured cognitive and affective mentalising abilities and the SST measured explicit and spontaneous empathic accuracy. No participants reported having previously read or heard of the SST stimulus text. Inter-rater

reliability for the SST was assessed using 10% of transcripts rated by the experimenter and two independent coders, with Fleiss's kappa calculated for spontaneous empathic inferencing, and Kendall's coefficient of concordance for explicit empathic accuracy. Inter-rater reliability was good for the spontaneous item (based on classifications of Cohen's kappa, see Altman, 1999), $\kappa = .72$, $p < .001$, 95% CI [0.36, 1.08]), for the explicit mental state attribution items, $W = .77$, $p = .013$, and very good for the comprehension items, $W = .94$, $p = .003$ (see Gibbons & Chakraborti, 2011). Internal consistency was adequate overall ($\alpha = .75$), and Cronbach's alpha was .70 for explicit mental state attribution and .53 for comprehension individually. These values were above those of .54 and .31 reported by Dodell-Feder et al. (2013), who anticipated that alpha would be negatively impacted by the range of first- and second-order questions.

7.8.1.2.4 Trait empathy. Trait empathic abilities were measured using the perspective-taking ($\alpha = .71$), empathic concern ($\alpha = .76$) and fantasy ($\alpha = .70$) IRI subscales. Fourteen irrelevant items (e.g., "I have no difficulty in cooking a complicated meal") were included across the scales to mask the aims of the study.

7.8.1.2.5 Fiction-exposure. The adapted version of Mar et al.'s (2006) Author Recognition Test-Revised (ART-R), used in Study 2 (Chapter 6), was used to facilitate measurement of both fiction- and nonfiction-exposure (Appendix E). The test comprised 55 fiction names, 55 nonfiction names, and 40 foils (150 items total). Internal consistency was good with this sample ($\alpha = .87$).

7.8.1.3 Procedure. Participants were randomly assigned to one of three story conditions: high cognitive load during reading, low cognitive load during reading, or no reading. First, the demographic questions and control measures (randomised within and between scales) were administered. Participants in the no reading group then proceeded to the next section of the experiment, and participants in the reading groups were asked to read the

story text while keeping in mind an eight-digit number (high cognitive load) or a two-digit

number (low cognitive load). After reading, the story participants were asked to report the

number they had been given, to answer basic questions about the story (attention check), and

to complete the narrative immersion scales (internally randomised and presented in a random

order). Participants were then asked to complete the Yoni test, followed by the SST, always in

this order to temporally separate the two reading sessions (Joyce stimulus text and

Hemingway SST text, Figure 12).
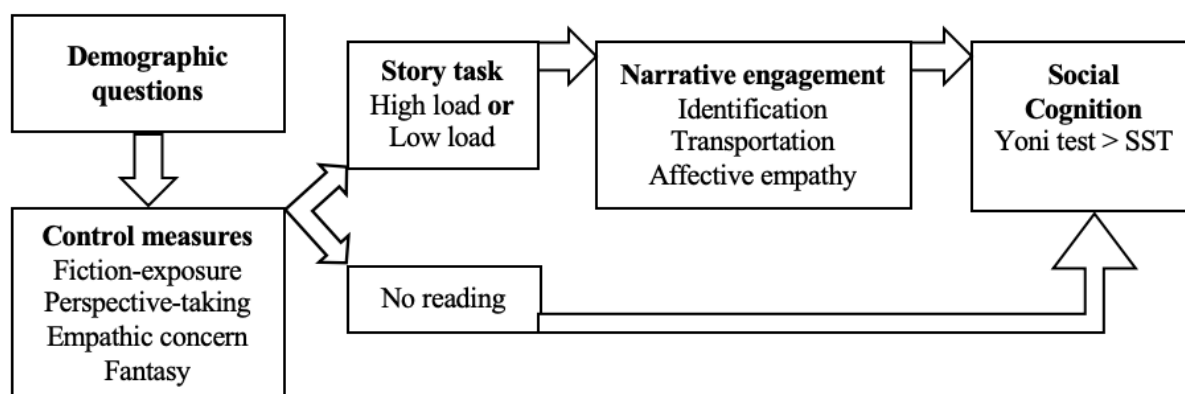
Figure 12. Flow diagram illustrating the experimental procedure. Note that participants in the reading groups read two stories (the Joyce stimulus text and the Hemingway text used in the SST) and participants in the no reading group read only the Hemingway story as part of the SST measure.

   7.8.1.4 Data analysis. As in the previous experiments, scale scores were averaged

across items in order to facilitate interpretation of their relationships to scale point labels.

Correlations were used to examine relationships among variables, and the effects of the

cognitive load manipulation were assessed using independent *t*-tests, and ANCOVAs when

covariates were included. Effects of reading group on Yoni test scores and SST explicit scores

were reported with and without the inclusion of covariates (using ANOVAs and ANCOVAs,

respectively), and the effect of reading group on SST spontaneous scores was assessed using

binary logistic regression. Due to the non-significant effect of the manipulation (reported

below), the analyses compared the three reading groups (high load, low load and no reading)

and reading versus no reading (high and low load groups combined). Confidence internals were bias-corrected and accelerated using bootstrapping ($N = 1000$).

7.8.2 Results

7.8.2.1 Group and gender differences. Trait fantasy, empathic concern and perspective-taking did not differ significantly between the reading conditions, all $Fs < 1.10$, all $ps > .33$, and neither did fiction-exposure or nonfiction-exposure, both Welch's $Fs < 3.00$, both $ps > .05$. There were no gender differences on any of the SST explicit or spontaneous dimensions, or on the Yoni test cognitive, affective or physical trials, all $ps > .54$. Females scored higher on empathic concern, $t(91) = -.28$, $p = .007$ [-3.67, -3.93], and lower on nonfiction-exposure, $t(91) = 3.25$, $p = .002$ [0.54, 4.08], compared to males.

7.8.2.2 Data distribution. Data were non-normally distributed across several scales (SST, ART-R and Yoni test dimensions, character identification and empathic concern). This appeared to result from the presence of outliers across all of the scales concerned (assessed using boxplots). These were retained in the analysis, and robust (Spearman) correlations were conducted.

7.8.2.3 Inter-scale correlations. Raw correlations between reading condition, immersion and empathic accuracy scores are presented in Table 12. Inter-scale correlations were in line with previous research showing associations among the subscales of the IRI (Davis, 1983), explicit and implicit empathic inferencing abilities in NT adults (Kanske et al., 2015), as well as fiction- and nonfiction-exposure (Fong et al., 2013; Mar et al., 2006). The immersion scales positively correlated with perspective-taking and empathic concern, and fantasy with transportation and character identification, but not with affective empathy. Yoni test cognitive and affective trials were associated with explicit empathic accuracy on the SST, and cognitive trials with spontaneous empathic inferencing on the SST, but neither measure was associated with the trait empathic ability (IRI) scales.

7.8.2.4 Effect of manipulation. Average immersion scores for each load group are presented in Figure 13. Transportation, identification and affective empathy scores were low compared to the text pretest in the high cognitive load group ($M = 3.70$, $SD = 0.83$; $M = 2.60$, $SD = 0.96$; and $M = 2.87$, $SD = 0.89$, respectively) and in the low cognitive load group ($M = 3.53$, $SD = 0.95$; $M = 2.54$, $SD = 0.90$; and $M = 3.01$, $SD = 0.92$, respectively). Mean differences were non-significant for all of the three dimensions (smallest non-significant $p = .457$ for transportation). Fantasy did not differ between the high and low cognitive load groups, t(61) = -.61, $p = .546$, 95% CI [-2.94, 1.57], and the effects of condition remained non-significant when fantasy was included as a covariate, all Fs < 1. Mean differences were non-significant for transportation, $t(61) = -.76$, $p = .457$, 95% CI [-.62, .28], identification, $t(61) = -.23$, $p = .816$, 95% CI [-.05, .23], and affective empathy, $t(6) = .60$, $p = .546$, 95% CI [-.32, .60]. There was no difference between the high and low cognitive load groups on fantasy, t(61) = -.61, $p = .546$, 95% CI [-2.94, 1.57], and the effects of condition remained non-significant when fantasy was included as a covariate, all reading condition Fs < 1. The assumption of homogeneity of variances was met for all tests.

The extent to which participants made an effort to retain the number in the high cognitive load condition may have attenuated the association between reading condition and immersion levels. To assess this, accuracy of recall was correlated with the immersion. All participants in both cognitive load conditions attempted to report the number they were asked to keep in mind. In the low load condition, all responses were accurate. In the high load condition, correct digits were given a score of 1, resulting in scores between 0 (0% accuracy) and 8 (100% accuracy). In cases where participants had reported fewer than eight digits, individual values were coded from left-to-right until a value was inaccurate and then from right-to-left in order to establish the location of the missing values (e.g., a recalled number of

318649 would receive a score of 6 against the true number of 31875649). There were no associations between recall and transportation, $r(26) = -.20$, $p = .32$, 95% CI [-.54, .15], character identification, $r(26) = -.14$, $p = .49$, 95% CI [-.47, .20], or affective empathy, $r(26) = -.09$, $p = .67$, 95% CI [-.37, .18]. This pattern held when trait fantasy was controlled, all rs < .07, all ps > .748.
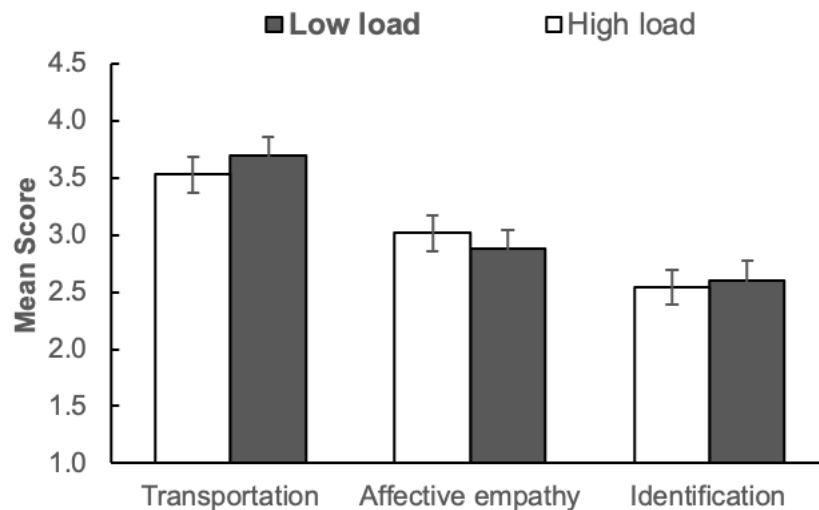
Figure 13. Raw transportation, affective empathy and identification scores presented for high and low cognitive load conditions (Experiment 5). Affective empathy and character identification were measured on 5-point Likert scales, and transportation was measured on a 7-point Likert-style scale. Error bars are standard errors of the mean.

7.8.2.5 Exploratory analysis: Effects of fiction-reading on empathic accuracy. As no differences in immersion were detected between the two reading (high and low load) groups, they were combined in order to enable comparisons between participants who read the stimulus story and those who did not. The independent variable for the following tests was reading, with reading ($n = 63$) versus no reading ($n = 30$) conditions. Nonfiction-exposure was higher in the reading versus the no reading group, $t(91) = -2.35$, $p = .024$, 95% CI = [-1.38, 1.19], but there were no group differences in fiction-exposure, or on the trait empathy scales, all ts < ± .67, all ps > .514.

Table 12

Correlations between Fiction-exposure, Trait Empathy, Narrative Immersion and Post-test Empathic Accuracy in Experiment 5.

	2	3	4	5	6	7	8	9	10	11	12	13
1. Fiction-exposure	.64*** [.49, .78]	.01 [-.17, .19]	.05 [-.15, .23]	.04 [-.14, .21]	.33** [.16, .53]	.30* [.07, .49]	.20 [-.08, .44]	.211 [-.003, .43]	.10 [-.09, .23]	.08 [-.14, .28]	.06 [-.14, .25]	-.09 [-.27, .11]
2. Nonfiction-exposure	-	-.03 [-.24, .18]	.03 [-.16, .22]	.02 [-.18, .20]	.19 [-.03, .39]	.13 [-.10, .36]	.17 [-.11, .43]	.12 [-.14, .36]	.17 [.01, .33]	.09 [-.07, .22]	.09 [-.11, .29]	-.11 [-.29, .08]
3. Fantasy		-	.34*** [.13, .54]	.45*** [.17, .67]	.20 [-.03, .40]	.39** [.19, .56]	.22 [-.04, .48]	.47*** [.25, .66]	.08 [-.19, .37]	.12 [-.19, .41]	.06 [-.20, .27]	.14 [-.13, .39]
4. Perspective-taking			-	.58*** [.45, .71]	.09 [-.13, .30]	.49*** [.28, .65]	.54*** [.34, .71]	.45*** [.27, .62]	.08 [-.15, .31]	.06 [-.17, .31]	.21 [-.02, .43]	.12 [-.12, .36]
5. Empathic concern				-	.05 [-.17, .28]	.42*** [.19, .62]	.50*** [.34, .64]	.49*** [.31, .63]	.14 [-.11, .39]	.13 [-.16, .40]	.09 [-.16, .31]	.11 [-.15, .35]
6. SST story comprehension					-	.17 [-.10, .38]	.08 [-.16, .32]	.09 [-.19, .33]	.17 [-.13, .47]	.16 [-.14, .44]	.39*** [.17, 57]	.13 [-.11, .40]
7. Transportation						-	.67*** [.49, .81]	.81*** [.71, .99]	.05 [-.19, .30]	.19 [-.08, .46]	.34** [.13, .53]	.16 [-.10, .43]
8. Affective empathy							-	.72*** [.53, .86]	.06 [-.17, .28]	.15 [-.11, .40]	.28* [.05, .47]	.07 [-.20, .32]
9. Character identification								-	.02 [-.20, .26]	.18 [-.06, .40]	.34** [.11, .55]	.16 [-.10, .39]
10. Yoni cognitive									-	.69*** [.53, .80]	.23* [.05, .42]	.22* [.02, .42]
11. Yoni affective										-	.39*** [.24, .55]	.17 [-.02, .36]
12. SST: Explicit											-	.29** [.10, .48]
13. SST: Spontaneous												-

Note. N = 93; * p < .05, ** p < .01, *** p < .001; 95% bias-corrected and accelerated confidence intervals are presented in brackets

7.8.2.5.1. Yoni test. Average Yoni test scores for participants who read and those who did not read are presented in Table 13. A mixed ANCOVA with reading condition as the between-participants variable, Yoni test trial type (cognitive, affective), and order (first, second) as within-participants variables, and first and second-order physical (control) scores as covariates, was used to assess the effect of reading on scores. Only two-way interactions between second-order physical trials (covariate) × trial type, $F(1, 89) = 7.15$, $p = .009$, $\eta_{p2} = .07$, second-order physical trials × order, $F(1, 89) = 35.65$, $p < .001$, $\eta_{p2} = .29$, and trial type × order, $F(1, 89) = 4.75$, $p = .032$, $\eta_{p2} = .05$, and the three-way interaction between these variables, $F(1, 89) = 8.67$, $p = .004$, $\eta_{p2} = .09$, were statistically significant. The main effect of reading and two-way interactions involving the reading condition were non-significant, all ps > .10, as was the three-way interaction, $F < 1$.[21] Additionally controlling for age, gender, fiction-exposure, and the three trait empathy scales revealed the same pattern of results (for effects involving reading condition, and this was the same with the inclusion of nonfiction-exposure, which had been found to differ between groups, all Fs < 1.11). These same tests were run with the high and low cognitive load conditions separated and the same pattern of results emerged with no significant effects of reading condition (for main and interaction effects involving reading condition, all Fs < 1).

7.8.2.5.2 SST. The effect of reading group (reading versus no reading) on SST explicit scores was assessed using *t*-tests and ANCOVA. There was a non-significant effect of reading group on SST explicit scores, $t(91) = .19$. $p = .851$, 95% CI [-1.19, 1.44], though controlling for trait empathy, age, gender, fiction-exposure and comprehension of the SST text, the effect reached statistical significance, $F(1, 84) = 4.00$, $p = .049$, $\eta_{p2} = .05$. Comprehension of the

---

[21] The data for the order 1 cognitive trials were negatively skewed, and this was not improved using square root or log transformations. Covariance matrices were unequal (assessed using Box's test, $p = .011$) and there was heterogeneity of variances in the cognitive order 1 data indicated by Levene's test ($p = .001$). However, ANOVA tends to be robust to such deviations (Norman, 2010) and so this was unlikely to substantially alter the result.

SST, the only other significant predictor, accounted for the majority of variance, $F(1, 84) =$ 25.87, $p < .001$, $\eta_{p2} = .24$. Addition of nonfiction-exposure, which was found to differ between the reading groups as reported above, did not alter the effect of reading condition, $F(1, 83) = 4.34$, $p = .040$, $\eta_{p2} = .05$. However, when the three reading conditions were separated out (high cognitive load, low cognitive load and no reading) the effect of condition was no longer statistically significant, $F(2, 83) = 2.77$, $p = .069$, $\eta_{p2} = .06$. This result was the same when nonfiction-exposure was included in the model, $F(2, 82) = 2.71$, $p = .073$, $\eta_{p2} =$ .06, and was likely due to the reduction in power associated with the decompression of the reading group. The assumption of homogeneity of regression slopes was met for all IV $\times$ covariate interactions, except for fiction-exposure in the model using the reading versus no reading IV ($p = .045$).

Binary logistic regression was used to predict the likelihood of making a spontaneous mental state inference on the SST using the same variables (reading versus no reading, age, gender, the three trait empathy scales, and fiction-exposure). The model was significant, $\chi_2(8)$ = 16.63, $p = .034$), and prediction success was 63% (71% for those who did not make a spontaneous empathic inference and 55% for those who did), which was an improvement compared to the baseline model (55%). Nagelkerke's R_2 of .22 indicated a moderately weak prediction. Only story comprehension, $p = .006$, and fiction-exposure, $p = .018$, significantly contributed to the prediction, such that a one-point increase in story comprehension was associated with an increase in the odds of making a spontaneous mental state inference of 1.34, and a one-point increase in fiction-exposure was associated with a decrease in odds (0.83). Reading condition was not associated with a change in the odds of making a spontaneous mental state attribution, $p = .119$.

When nonfiction-exposure, which differed between the reading versus no reading groups (as reported above), was included in the model, fiction-exposure no longer accounted

for a change in odds, p = .116. In this model, $\chi_2(9)$ = 17.54, p = .041, Nagelkerke's R_2 = .23, comprehension of the SST story was the only significant predictor, p = .006, and accounted for an increase in the odds of making a spontaneous mental state inference of 1.34 for each one-point increase in comprehension. Using the same variables and separating out the cognitive load conditions (high load, low load and no reading), revealed the same pattern of results.

Table 13

Mean Yoni Test Scores (with Standard Deviations) for Reading and No Reading Groups.

Trial Type	Order	Reading	No reading	Total
Cognitive	1	11.49 (0.91)	11.80 (0.48)	11.59 (0.81)
	2	19.51 (3.90)	18.80 (4.24)	19.27 (4.00)
Affective	1	11.25 (1.26)	11.43 (0.90)	11.31 (1.15)
	2	29.54 (4.99)	29.57 (4.68)	29.55 (4.87)
Physical	1	7.43 (1.30)	7.43 (1.33)	7.43 (1.35)
	2	5.03 (1.51)	5.33 (1.45)	5.13 (0.15)

7.8.3 Discussion

The aim of the present study was to examine the effects of narrative immersion on empathic accuracy task performance, taking the multidimensionality of immersion and empathic accuracy into account. Following a pilot study in which an attempt to manipulate narrative immersion levels failed, a pretest identified an immersive stimulus text, and two pilots established a method for varying immersion, which led to the employment of a cognitive load manipulation in the final experiment. In that final experiment, however, the cognitive load manipulation failed to impact immersion levels: immersion did not differ across the high and low cognitive load conditions. This may have been due to a combination of factors including low immersion overall, and low power (a post-hoc power analysis was conducted, which indicated that the study had very low power to detect the effect of the

experimental manipulation). Consequently, Hypotheses 12-15, which predicted effects of immersion on empathic accuracy domains, were untested. Instead, correlations between immersion and story comprehension, empathic accuracy, trait empathy, and fiction- and nonfiction-exposure were identified, and the effect of reading versus no reading on empathic accuracy was assessed in order to test the general prediction that fiction-reading would increase empathic accuracy.

7.8.3.1 Immersion and empathic abilities. Scales within measures tended to be positively correlated, and this was line with previous research (Davis, 1983; Fong et al., 2013; Kanske et al., 2015; Mar et al., 2006). There was a lack of correlation between self-report and behavioural empathy measures which also reflected previous findings, and the multidimensionality of empathy (e.g., Ickes, 1997; Mar et al., 2006): neither the Yoni test nor the SST were associated with self-report concern, perspective-taking or fantasy in either the first pilot experiment (Experiment 1 in this chapter) or the revised experiment (Experiment 5 in this chapter). The self-report scales require people to reflect on their tendencies to feel concern for others, to take others' points of view or to become immersed in narratives, rather than the ability to *accurately* interpret others' inner states, which is the subject of the behavioural tasks. Moreover, Yoni test and SST scores did not correlate. Differences in the stimuli (the mental states of schematic faces versus relatively naturalistic story characters) may entail different inferencing strategies. The extent that mentalising versus experience-sharing strategies are employed (see Chapter 2) may both influence empathic accuracy and be influenced by immersion. The lack of correlation among measures emphasises the complexity of the empathy construct and the need for a multidimensional approach to measurement.

The immersion scales, transportation, character identification and affective empathy, all positively correlated with trait perspective-taking and empathic concern. This suggests that similar processes may underpin the ways that people think about both fictional characters and

other people in their daily lives. Furthermore, character identification and transportation were positively associated with fantasy; however, affective empathy was not. The Empathy Index contains items such as "sympathetic" and "warm", which reflect the concern component, rather than the imaginative abilities probed by items in the identification and transportation scales, highlighting the multidimensionality of narrative immersion (de Graaf et al., 2012).

The immersion scales were all positively associated with explicit mental state reasoning on the SST but not with the Yoni test. In other words, becoming immersed in one stimulus story appears to relate to the ability to accurately interpret the mental states of characters in another story. It may be that immersion supports the interpretation of characters depicted in fiction, but that this effect does not generalise to other targets. Alternatively, immersion may support the empathic inferencing processes probed by the SST but not those probed by the Yoni test. For example, immersion in stories may particularly facilitate experience-sharing rather than cognitive mentalising (e.g., Stansfield & Bunce, 2014), and experience-sharing may have been induced by the SST text. While the direction of cause remains uncertain, these findings support the proposition that dimensions of immersion and empathic abilities are related.

7.8.3.2 Fiction-reading and empathy.

7.8.3.2.1 Fiction-exposure. As anticipated based on prior research (Mar et al., 2006, 2009), fiction-exposure positively correlated with nonfiction-exposure in the present study, and with transportation and story comprehension on the SST. Unexpectedly, it was not associated with trait empathic abilities in this sample. As the size of the correlation between fiction-exposure and trait empathy is small in magnitude (Mumper & Gerrig, 2017), the present study may have lacked sufficient power to detect it. Fiction-exposure increased the odds of making a spontaneous mental state inference on the SST, but not when nonfiction-exposure was included in the model. This indicates that the increased likelihood of reflecting

on the mental states of characters was not accounted for by fiction-exposure alone. Rather, both fiction and nonfiction may explain some variance, but not enough to be individually significant (which may partly reflect a reduction in precision associated with the increase of parameters in the latter model).

7.8.3.2.2 Fiction-reading versus no reading. Chapter 3 reviewed research that has cast doubt on the assertion that fiction-reading, compared to nonfiction-reading or no reading, enhances empathic accuracy task performance (e.g., Camerer et al., 2018; Panero et al., 2016, Samur et al., 2018). In the present exploratory analysis, the effect of fiction-reading on explicit SST scores reached significance, but only when holding constant trait empathic abilities, age, gender, comprehension of the SST text and ART scores, $p = .049$, and $p = .040$, when nonfiction-exposure was controlled). This provides some support for the causal impact of reading on empathic accuracy.

The stimulus text used in the present study was a short passage extracted from a longer story. Causal effects of fiction may develop through exposure to longer passages or via repeated immersive experiences, whereas short experimental manipulations may not be sufficient to detect effects (Mar, 2018b). Indeed, the effects of narrative persuasion appear to be progressively internalised over a period of incubation ("sleeper effect"; Appel & Richter, 2007; Kumkale & Albarracín, 2004), and the same may be true of effects on empathic accuracy. Playing prosocial video games has been shown to lead to subsequent prosocial behaviour (Gentile et al., 2009) and high transportation to increase dispositional empathy two weeks after reading fiction (Bal & Veltkamp, 2013). Therefore, the effects of immersion on empathic accuracy may not be observed through immediate testing, instead requiring a longitudinal approach (e.g., Pino & Mazza, 2016). Alternatively, immediate effects may occur, but they may be small in magnitude, necessitating high-powered studies (see Black & Barnes, 2015b, for an example of causal effects shown using a within-subjects design). This

may go some way to explain the mixed literature on fiction effects on empathic accuracy while the correlational research seems robust. For example, De Mulder et al. (2017) found no direct effects of reading literary texts on Yoni test and RMET performance, though they did find an effect of lifetime exposure to fiction on Yoni test cognitive trials. Further research is needed to establish the direction, induction and timespan of causal effects (see also Mar, 2018b; Mar, Oatley, Djikic & Mullin, 2011).

7.8.3.3 Limitations. The attempt to manipulate immersion failed across all three immersion dimensions and so the primary hypothesis that high immersion would lead to higher empathic accuracy compared to low immersion could not be tested. It is possible that the present study failed to detect the effect of the manipulation. Despite a pretest establishing a relatively immersive text (albeit with low power), and pilot studies identifying a procedure for manipulating immersion levels, they were low across the sample. The stimulus text passage may not have sufficiently induced immersion, and this may have attenuated the effect of external cognitive load instructions (see Green & Brock, 2000, Experiment 4). The absence of a no-load control condition (rather there was a no reading control condition) means that these possibilities could not be explored using the present dataset, and future research may usefully address this limitation.

Sample characteristics may have influenced the results of this study. First, although age and gender were controlled in Experiment 5, most of the participants were female, which limits generalisability of the results. This is particularly relevant as gender differences have been shown in empathic abilities (e.g., Davis, 1980). Second, the pretest and manipulation pilot participants volunteered online in return for payment, whereas the participants for the main experiment were recruited via a university research participation scheme in return for course credit. As a result, motivational or situational factors may have accounted for differences in the ability to become immersed: although instructions were presented using the

same online platform for both lab and online participants, lab participants attended advance-booked appointments whereas online pretest and pilot participants may have participated at a time when they felt motivated and distractions were minimal. Research has indicated that online participants may play closer attention to manipulation instructions than traditional subject pool participants (Hauser & Schwarz, 2016). The potential impact of sample characteristics may be explored in relation to this field of research in the future as it raises a broader question about the generalisability of fiction-engagement effects across populations.

7.8.3.4 Contribution and future directions. The present study did not support the hypothesis that immersion would increase empathic accuracy, neither did it refute the proposition. The correlation analysis revealed positive associations between immersion, dimensions of trait empathy and explicit empathic accuracy (effects were medium-large; see Cohen, 1988). Specifically, transportation and identification with character were positively associated with the trait empathy scales, perspective-taking, empathic concern and fantasy. The third immersion scale, which measured empathic feelings during reading, was associated with perspective-taking and empathic concern. Furthermore, all three immersion scales showed positive associations with a behavioural measure of empathic accuracy: the explicit mental state attribution questions on the SST (effects were small-medium; see Cohen, 1988). These findings support the view that immersion in literature and empathic abilities are associated. It may be that immersion does play a role in fiction effects on empathic accuracy as it has been shown to do in the field of narrative persuasion, but that the experiment reported here did not detect the causal effect (due to the insufficient manipulation and low overall power). Alternatively, people who are more empathic may tend to become more immersed in fiction. Future research addressing the limitations of the current study may establish the direction of this relationship.

The potential relationship between immersion in fiction and subsequent empathic accuracy task performance may be of relevance to researchers from other fields who employ narrative-based tools to test empathy levels (e.g., the MASC and SST; see Chapter 2). Immersion in the stimulus text for the SST was shown to be low on average in the present study; however, immersion varies across individuals and may be differently induced by fictional content.[22] Future research may help to clarify the role played by individual differences in immersion—and its antecedents—in fiction effects on empathic accuracy domains. This, in turn, could have implications for researchers using narratives as empathy task stimuli, who might consider accounting for the role of immersion in their research designs.

The present study lent some support to the general causal hypothesis that fiction-reading immediately enhances empathic accuracy: participants who read the popular fiction passage compared to those who did not, scored higher on the SST when individual differences in trait empathy, fiction habits, comprehension and demographic differences were controlled. However, this test measured empathic accuracy for story characters (albeit from a second story) and so it is unclear whether the effect would generalise to other contexts (there was no effect on the Yoni test empathic accuracy measure). Furthermore, the effect was only demonstrated when controlling for a range of individual differences and so further research is needed to establish the stability of this effect. In view of the mixed results reported using experimental designs (e.g., Camerer et al., 2018; Panero et al., 2016; Samur et al., 2018; van Kujik et al., 2018), further research may continue to examine the influence of individual differences. This could facilitate the development of a comprehensive model of fiction effects

---

[22] In the present study, one participant experienced a strong emotional reaction to the Joyce story that was incongruent with the characters' depicted emotional levels, indicating a distinctly personal (non-isomorphic) connection with the characters' circumstances. This illustrates the relevance of individual–perhaps autobiographical (e.g., Bluck, Baron, Ainsworth, Gesselman & Gold, 2013)–factors in affective engagement with story characters.

on empathic accuracy, including mediators and moderators (for example, see Mar's, 2018a, Social Processes and Content Entrained by Narrative model, which maps the pathway from story exposure to enhanced social cognition, and acknowledges moderators such as immersion, personality, and situational factors).

The present study emphasised the need for a multidimensional approach to immersion, as well as to fiction and empathy in order to clarify the nature of their relationships. Future research could explore a range of texts in order to produce a more granular perspective on the textual features that enhance empathic abilities and the cognitive process through which such effects occur. Clarifying the scope of fiction as a tool for enhancing social skills could have implications for work with groups that experience difficulties in interpreting people's thoughts and feelings (e.g., ASD), as well as downstream consequences such as increasing altruistic behaviours (e.g., Johnson, 2012).

7.9 Summary and Moving Forward

Fiction-readers "not only enter a narrative world, they may also become highly involved with the people they find there" (Green & Brock, 2000, p. 702). Fiction enables readers to experience a range of events in the lives of human or humanlike characters (Mar & Oatley, 2008) and so fiction-reading may benefit real-world empathic accuracy. Several studies have reported relationships between lifetime exposure to fiction and empathic abilities, and others have shown small, positive effects of fiction on empathic accuracy after reading. However, the mechanisms associated with such effects remain unidentified. Through a series of experiments, the present study established a procedure for manipulating levels of immersion in a popular fiction text. In the final experiment, however, the manipulation failed to distinguish group immersion levels, and so causal inferences about the impact of immersion could not be made. A correlation analysis showed positive associations between immersion dimensions and both self-report (perspective-taking, empathic concern and fantasy) and

behavioural (interpreting the mental states of story character) empathic ability measures. An exploratory analysis revealed a positive effect of reading compared to no reading on empathic accuracy, though further research is required to establish the stability of this effect. Future research employing high-powered and longitudinal designs may elucidate the antecedents and consequences of immersion in fiction, which may help to develop models of fiction effects on empathic abilities and subsequent prosocial behaviours, and have implications for real-word social interventions.

The present study concludes the empirical work undertaken in the current project. Two previous studies, presented in Chapters 5 and 6, examined relationships between fiction-exposure and empathic abilities. The present study aimed to contribute to the causal literature through an investigation of the effects of narrative immersion on empathic accuracy. Due to the failure to manipulate immersion levels, causal inferences about the effects of immersion could not be made, though there was limited support for the causal hypothesis that reading fiction enhances empathy compared to not reading. Correlations between immersion dimensions and empathic abilities provided evidence for the assertion that immersion in narratives and empathic abilities are associated. The next and final chapter in this thesis aims to summarise and synthesise the results from all three studies, discuss their implications in the context of the theoretical framework established in Chapters 2 and 3, identify the strengths and limitations of the research project, and highlight avenues for future enquiry.

Chapter 8: General Discussion

The previous three chapters reported the studies that constitute the empirical work of this thesis. This chapter will recap the main research aim, and discuss the findings in relation to extant literature and the multidimensional approach proposed in Chapters 2 and 3. Factors contributing to relationships between fiction and empathy will be addressed, methodological limitations identified, and the extent to which the findings support the theory that fiction benefits empathy will be discussed. The assumptions that fiction is beneficial for empathy, and that empathy itself is beneficial, will be critiqued. The chapter will conclude by establishing the original contribution and potential implications of the present research and consider avenues for further study.

8.1 Recap of Main Aim

> I have often noticed that we are inclined to endow our friends with the stability of the type that literary characters acquire in the reader's mind. No matter how many times we reopen "King Lear," never shall we find the good king banging his tankard in high revelry, all woes forgotten, at a jolly reunion with all three daughters and their lapdogs. Never will Emma rally, revived by the sympathetic salts in Flaubert's father's timely tear. Whatever evolution this or that popular character has gone through between the book covers, his fate is fixed in our minds, and, similarly, we expect our friends to follow this or that logical and conventional pattern we have fixed for them. (Nabokov, 1955/2000, p. 265)

Bewildered at the incongruity of an old friend's behaviour, Nabokov's narrator explains that he expects real people to show consistency, just as fictional people tend to do. Telling and receiving stories is a practice so ingrained (MacDorman, 2019) that people commonly use fictional characters and their interactions as referents for real-world people and *their* interactions. Both are expected to exhibit coherence and behave, more or less, predictably. In Nabokov's passage, the narrator is, himself, fictional, as is the friend that he refers to, invoking Oatley's (1999) proposition that fiction can constitute a valid imitation of the social

world. Oatley, Mar and colleagues have argued that by providing a simulation of the social

world, fiction can recruit and enhance the cognitive mechanisms associated with

understanding and responding to real-world others (Mar & Oatley, 2008; Oatley, 1999;

2011b; Oatley & Djikic, 2017).

In the real world and in fiction, prior knowledge of a person is often sufficient to

interpret their thoughts or feelings in a given situation. Sometimes, however, that knowledge

is limited or non-existent, and so the perceiver must base their inferences on other knowledge,

facial expressions, gestures or verbalisations. Alternatively, they can generate inferences

based on how they, themselves, would respond to the same situation, and generalise those

feelings to the other person. These two approaches represent mentalising and experience-

sharing strategies for empathic accuracy: the measurable ability to accurately infer thoughts

and feelings. Consequently, a prosocial concern component—the desire to help, or to alleviate

the other person's suffering—may be activated, which can, in turn, initiate prosocial

behaviour. Empathy, therefore, is multidimensional (e.g., Davis, 1980): a person with strong

mentalising skills does not necessarily experience high levels of concern, and correctly

interpreting emotions may not equate to correctly determining intentions (e.g., Oakley et al.,

2016). Fiction is multidimensional too: stories vary along theme and style dimensions, can be

presented via different media channels and induce different immersive experiences. The

present research aimed to examine links between modes of narrative engagement and

empathic processes, in order to provide a granular perspective on the relationship between

fiction and empathy.

8.2 Current Findings

The general premise that fiction enhances empathic abilities was explored in the current

research using a multidimensional approach. Mentalising and experience-sharing strategies,

empathic concern and prosocial behaviour components were measured using self-report tools

requiring participants to reflect on their skills and behaviours (Study 1), and behavioural measures to assess empathic accuracy for cognitive and affective content (Studies 2 and 3). Relationships between empathic abilities and fiction-exposure were assessed (Studies 1 and 2), including exposure to fiction via different media formats (Study 1). The causal effects of becoming immersed in a passage of fiction were investigated (Study 3). Overall, the findings lent some support to the theory that fiction enhances empathic abilities, although the most promising evidence was correlational.

8.2.1 Fiction is multidimensional: Summary and discussion of Study 1 results

If stories themselves are universal, the way we tell them changes with the technology at hand. Every new medium has given rise to a new form of narrative. (Rose, 2011, p. 2)

I find television very educating. Every time somebody turns on the set, I go into the other room and read a book. (Attributed to Groucho Marx; The Reader's Digest Association, 1950)

Study 1 aimed to examine the extent to which self-report empathic abilities relate to lifetime fiction-exposure via different media formats. Additionally, it aimed to explore how far people's preferences for media, genre, and acting experience would be associated with components of empathy.

8.2.1.1 Media. A new fiction-exposure tool, the FMET, was created to test how far fiction-exposure via prose, film and plays would predict variation in empathic abilities. The FMET was found to positively predict variation in self-report altruistic behaviours and self-report fantasy, but not empathic concern or perspective-taking, supporting the general prediction that fiction-exposure is differentially associated with dimensions of empathy. The positive prediction for altruism contributes to limited previous research identifying positive effects of fiction on prosocial behaviour (e.g., Johnson, 2012). The positive prediction for fantasy (which has been identified as an experience-sharing measure; Chapter 2, Section 2.5), also aligns with previous research showing that experience-sharing is the component most

impacted by fiction (when previous studies' effect sizes were averaged within empathy dimensions; Chapter 3, Section 3.5.2). Results for perspective-taking and empathic concern, however, diverged from previous findings: Mumper and Gerrig's (2017) meta-analysis identified positive associations between fiction-exposure and both scales. In the current study, positive predictions for both were statistically significant but the FMET did not significantly contribute to either prediction; rather, gender was the only contributor. Therefore, self-report mentalising and concern, as measured using the IRI subscales, appear to be impacted by gender, and perhaps by other, unmeasured individual differences, which will be addressed later in this chapter.

At the level of individual associations, print and play-exposure positively correlated with altruism, and print-exposure with fantasy (all fiction scales positively correlated with fantasy when age, gender, education and English language fluency were controlled, highlighting the influence of individual difference variables). In line with Mumper and Gerrig's (2017) meta-analysis, effects were small in magnitude. Relationships with empathic dimensions were similar for people's preferred media formats. Additionally, television was included in the preferences scales, and it negatively correlated with empathic concern, perspective-taking and altruism. Preference for film also negatively correlated with altruism (though film-exposure showed a positive relationship with altruism). These findings could reflect variation across narrative structure (story arcs in TV drama serials are distributed across several episodes) or engagement with exposition via news, documentary, and reality TV, rather than with fictional narratives. However, a follow-up analysis revealed no association between preference for TV and enjoyment of nonfiction, though there was a negative association with enjoyment of postmodern narratives. Future research could usefully examine the combined effects of media presentation and thematic content.

8.2.1.2 Active engagement with medium. Visual cues provided in film and TV may alter the extent that mentalising and experience-sharing processes are activated and exercised. Film is "plagued" by realistic effects, and because its narrative mode is "showing", it can be considered an expository medium (Bal, 1985/2009, p. 44). For example, film icon Bette Davis could reportedly dilate her pupils at will in order to highlight moments of arousal, aggression or ardour (Higham, 1982). Camera angles, close-ups and music can emphasise pivotal plot-points, generate atmosphere, and support characterisation. In contrast, readers must imaginatively conjure characters and events, entailing a more active and imaginative mode of engagement (see Hakemulder, 2000).

Narratives can be actively participated in via video games, roleplay and acting. Acting, in particular, requires the construction of character from inside that character (see Hakemulder, 2000) as actors must have sufficient knowledge of the character's experiences, beliefs and intentions to engender a convincing portrayal. In the present study, experience of acting was positively associated with altruism, although associations did not reach significance for the three IRI scales. It is not clear why acting experience was only associated with altruism; it may because the self-report acting experience scale did not reflect acting proficiency, only experience. Previous research has found that actors tended to score higher than the general population on self-report empathy scales (Nettle, 2006) and that acting classes may have a positive effect on mentalising and experience-sharing (e.g., Goldstein & Winner, 2012). These findings indicate that the empathy domains could be associated with acting but that effects were not detected in the current study. Again, this may reflect the influence of other, unmeasured variables (e.g., social and motivational factors, discussed later in this chapter).

8.2.1.3 Genre. Study 1 showed divergent relationships between enjoyment of different genres and empathic abilities, supporting the proposal that narrative modes have different

effects on empathy. In Mumper & Gerrig's (2017) meta-analysis, nonfiction-exposure was associated with perspective-taking and empathic concern, but associations were stronger for fiction. Similarly, in Study 1, enjoyment of nonfiction was associated with perspective-taking, empathic concern and fantasy, but associations tended to be larger for fiction (e.g., people who enjoyed drama and romance narratives were more likely to show higher levels of perspective-taking, empathic concern and fantasy than people that enjoyed nonfiction).

　　8.2.1.4 Active engagement with content. As anticipated, different fiction genres were associated with different empathic abilities: crime/thriller was not associated with any empathic abilities measured, whereas postmodern and comedy narratives both positively correlated with all four. This raises the question of what postmodern and comedy narratives have in common. Postmodern narratives contain features like unreliable narrators, fragmentation and multiperspectivity (the ability to track multiple perspectives; Nünning, 2014), and so the narrative arc may be complex, non-linear and its outcomes unexpected. Similarly, comic narratives can be incongruous and surprising (for a brief history of the "incongruity theory of comedy", see Morreall, 2009). Like reading, compared to other forms of media engagement (described above), these elements may invite an active style of engagement, which recruits the mechanisms associated with interpreting real-world others. On the other hand, comic characters and narratives can be formulaic and predictable, and they are often associated with commercial, rather than with literary media. Sitcoms, for example, are among the oldest and most ubiquitous forms of TV programming (e.g., Dalton & Linder, 2005). Nuanced differences within genres may be examined through future research aiming to identify the specific narrative ingredients responsible for the observed effects. In this way, future studies may extend the work of researchers like Koopman, Hakemulder, Van Peer and others (e.g., Hakemulder, 2004; Koopman, 2016; Van Peer et al., 2007) who have already

facilitated a more granular perspective of fiction effects through their explorations of specific textual features.

8.2.1.5 Multidimensionality. In Chapter 3 it was proposed that different modes of engagement with fiction may relate to different components of empathy, and the findings from Study 1 supported this proposition. Positive relationships between fiction and empathic abilities appear not to be unique to literary narratives, nor to fiction-in-print. Rather, medium and genre were both important in associations between fiction and empathic abilities, and narrative modes that invite an active, imaginative form of engagement may particularly support empathy. Based on the current findings, if required to choose a mode most likely to benefit empathy, printed prose would represent the best candidate. However, the print-exposure scale was based on established ART measures, whereas the film and play scales were designed for the study and may be developed and validated in the future.

The Study 1 findings contribute to previous research in three key ways: first, they highlight the importance of different media and genre formats in relationships between fiction and empathic abilities, as well as other factors that may influence these relationships which will be addressed later in this chapter. Second, they lend some support to the proposal that fiction particularly supports the experience-sharing inferencing strategy, due to the positive prediction for fantasy. Third, they demonstrate associations with altruism, which has received limited prior attention (only one prosocial behaviour effect size was included in Dodell-Feder & Tamir's, 2018, meta-analysis). Considering Study 1's positive prediction for altruism but not empathic concern (which can represent a precursor to altruism; see Chapter 2, Section 2.2), future research may examine pathways from fiction to prosocial behaviour that preclude the concern mediator.

8.2.2 Empathy is multidimensional: Summary and discussion of Study 2 results

If you can learn a simple trick, Scout, you'll get along a lot better with all kinds of folks.

You never really understand a person until you consider things from his point of view […]

until you climb inside of his skin and walk around in it. (Lee, 1960, p. 35)

Lee's character, Atticus Finch, advocates sharing in the experiences of others in order to understand and relate to them. Experience-sharing represents one approach to inferring others' mental states, and in Chapter 3 it was argued that regular fiction-readers may particularly excel at using this strategy (and Study 1's correlational evidence lent some support to this proposition). Study 2 aimed to evaluate the effectiveness of experience-sharing compared to mentalising for interpreting the mental states of others, and their relationship to fiction-exposure. Whereas Study 1 employed self-report measures of empathic abilities, Study 2 used a behavioural measure involving the attribution of affective states, in order to rule out the effects of socially desirable responding.

8.2.2.1 Contrasting empathic inferencing strategies. Study 2 participants were asked to estimate the emotional ratings given in response to a series of pictures by a group of target "experiencers". In one condition, participants viewed the same pictures as the experiencers, and were therefore able to base their estimates on their own emotional responses to the pictures. In the other condition, participants viewed pre-recorded videos of the experiencers viewing the pictures and were able to base their estimates on the experiencers' facial expressions. These conditions were mapped onto the mentalising and experience-sharing routes to empathic accuracy outlined in the empathy model established in Chapter 2. In the original study on which this task was based (Zhou et al., 2017), the experience-sharing route was found to be more effective than mentalising but was nevertheless undervalued by participants. As readers engage their own emotions as proxy for those of fictional others, it was hypothesised that fiction-exposure would lead to higher empathic accuracy overall, that

this effect would be greater when using the experience-sharing strategy, and that frequent fiction-readers would be more likely to appreciate the value of this strategy.

8.2.2.2 The impact of context. As discussed in Chapter 2, Ickes (1997) distinguished empathic accuracy from empathic inferencing: empathic accuracy represents a state of understanding attained using a given inferencing strategy. Intrinsic to any assessment of empathic accuracy, therefore, is the indexing of levels of error. The results of Study 2 showed that empathic accuracy differed as a function of mentalising versus experience-sharing strategy, as well as valence and target. In contrast to Zhou et al.'s (2017) findings, neither strategy was found to be more effective. The strategies entailed similar levels of error; however, this error was in opposite directions: estimates using the mentalising approach tended to be over-positive and estimates using the experience-sharing approach tended to be over-negative.

There are at least three reasons that this finding may have occurred. First, it could represent a systematic difference in mentalising and experiencing-sharing outcomes: the well-documented bias toward negative stimuli (Baumeister et al., 2001; Rozin & Royzman, 2001) operates when interpreting another person's facial expressions (see Kauschke et al.'s, 2019, review) but not when instantiating another's emotion, because people tend to regulate their own negative emotions in order to minimise negative states (Baumeister et al., 2001). Second, these differences could simply have emerged as an artefact of the contrasting stimuli used (pictures versus videos). For example, Johnson (2012) found that bias towards fearful facial expressions was only observed when faces were presented for longer (2 seconds compared to 500 milliseconds), indicating that deliberative rather than automatic processes facilitated the bias. Negative bias in the current study could have arisen as a result of attention paid to unfolding dynamic facial expressions compared to still images. Third, because the measure relies on the experiencers accurately interpreting and reporting their own emotional reactions

to pictures, original target ratings may have been subject to bias. To test whether the finding reflects a genuinely systematic effect of strategy, which represents a promising area of enquiry, future research could incorporate established or independently coded stimuli and attempt to vary inferencing strategy without varying the stimulus presentation mode.

8.2.2.3 The impact of fiction-exposure. Fiction-exposure and nonfiction-exposure were both positively associated with empathic accuracy, such that the more fiction or nonfiction people had been exposed to, the smaller their errors when estimating the experiencers' emotional ratings. However, relationships were stronger for fiction-exposure, and this was in line with the results from Study 1, where general preferences for fiction genres were more strongly associated with empathic abilities than preferences for nonfiction. It also aligned with previous research showing larger correlations between fiction-exposure and empathic abilities (Mumper & Gerrig, 2017). Furthermore, fiction-exposure was found to positively predict empathic accuracy when controlling for nonfiction-exposure (as well as when controlling for age and foil selection on the ART), and this effect held across both mentalising and experience-sharing strategy conditions.

The Study 2 findings supported the general prediction that higher fiction-exposure would be associated with higher empathic accuracy, but they did not support the hypothesis that this effect would be greater for participants in the experience-sharing condition. Similarly, fiction-exposure did not contribute to the likelihood of selecting one strategy condition over the other (an indirect measure of perceived value of the strategies). This suggests that both inferencing strategies may be enhanced by repeated exposure to fiction.

8.2.3 Effects of immersion in fiction on empathic accuracy: Summary and discussion of Study 3 results

A great book should leave you with many experiences and slightly exhausted at the end. You live several lives while reading it. (Styron interviewed in Cowley, 1958, p. 274).

Narrative invites getting "lost" (Nell, 1988, p. 8) or becoming immersed in a story, whereas exposition invites action (Keen, 2007). Study 3 examined the causal effects of immersion in fiction on empathic accuracy for cognitive and affective content using two established empathy measures (a naturalistic fiction-based task and a classic vignette task). Despite an initial pilot study, a text pretest, and two immersion manipulation pilots, the cognitive load manipulation ultimately failed to significantly vary any of the facets of immersion measured: transportation, identification with character, or feelings during reading. The effect of the cognitive load manipulation in Experiment 4 (manipulation pilot) was promising but was only observed in the small sample when trait fantasy was controlled.

It may be that cognitive load manipulations do not influence immersion when reading, because rather than requiring any cognitive elaboration, the effects of immersion may be peripheral or indirect (see Djikic & Oatley, 2014) and therefore unimpacted by cognitive load (e.g., Rocklage et al., 2018; Tormala & Petty, 2004; Zemborain & Johar, 2007). Furthermore, as social processes are practiced, they become less effortful and less likely to be impacted by competing mental tasks (Mar, 2018a), and so experience with getting immersed in stories may reduce the impact of the dual task paradigm. Future research could re-examine the utility of this method by additionally employing a no load control group, or by developing alternative approaches to manipulating immersion dimensions, and examining the effects on empathic accuracy post-reading. Doing so could shed new light on the pathways between immersion dimensions and empathic components.

8.2.3.1 Immersion and empathy. In Study 3, all immersion scales (transportation, character identification and affective empathy) were associated with the ability to accurately infer the mental states of story characters (SST explicit scores). Furthermore, the three immersion dimensions were generally positively associated with trait empathic abilities, which lent support to the overlap between empathy and immersion postulated in Chapter 3.

However, there was no correlation between affective empathy (the production of feelings during reading) and fantasy (which measures the tendency to engage imaginative processes as well as emotion-sharing). This highlights the multidimensionality of immersion, and that the cognitive mechanisms that serve certain features of immersion do not necessarily support empathic skills. The notion that fiction-engagement does not always benefit empathy will be discussed later in this chapter.

8.2.3.2 Causal effects of reading. An exploratory analysis revealed some supporting evidence for causal effects of reading a passage of fiction on empathic accuracy compared to not reading: reading appeared to increase empathic accuracy for story characters' mental states (SST explicit scores). Because there was no expository nonfiction-reading control group—the comparison was between fiction-reading and no reading—it is not clear how much variance can be attributed to reading the specific stimulus story versus general reading processes (although research showing effects of fiction and nonfiction, e.g., Mumper & Gerrig, 2017, suggests that some variance in empathic abilities is unique to fiction).

There was no causal effect of reading on the Yoni test measure of mentalising. Power may have been insufficient to detect the small effect of fiction-reading (Dodell-Feder & Tamir, 2018) on the Yoni task. Another explanation is that as the SST is a naturalistic measure that may recruit experience-sharing strategies, it captured variation in ability that the Yoni test did not probe. Alternatively, the reading exercise may have primed general comprehension of the SST text which, in turn, increased empathic accuracy scores (comprehension accounted for the majority of the increase in empathic accuracy). Comprehension may be essential to fiction effects, but it remains unclear whether the effects of comprehending one narrative correspond to increased accuracy for other (narrative or non-narrative) empathic content. Furthermore, as the SST may recruit both mentalising and experience-sharing processes, it is unclear how far each inferencing pathway contributed to

accuracy. The effect of fiction-reading on the SST was only observed when controlling for age, gender, fiction-exposure, and text comprehension, which emphasises the influence of other factors in fiction effects. Overall, evidence for the effect of fiction-reading on empathic skills was promising, but limited.

8.3 Factors Impacting Relationships between Fiction and Empathic Abilities

8.3.1 Individual differences

Previous research has shown that age and gender relate to empathic abilities and can interact with task stimuli (e.g., Richter, Dietzel & Kunzmann, 2010), that gender impacts fiction effects on empathic abilities (e.g., Mar et al., 2009), and that age positively predicts ART scores (e.g., Kidd & Castano, 2017a). In the present research, gender differences were shown on the IRI subscale measures of empathy, in line with Davis (1980, 1983), but not on the self-report altruism scale, which matched Rushton et al.'s (1981) findings. Males and females did not differ on immersion (which reflected, for example, Green & Brock, 2000, Experiments 1 and 4) or on the behavioural measures of empathic accuracy (in line with Dodell-Feder et al., 2013; in Study 3 Experiment 1, females scored higher than males on the SST, but this was not replicated in Experiment 5). Generally, exposure to fictional print did not vary by gender (which aligns with Mar et al., 2006). However, males recognised more film names than females on the FMET in Study 1 and more nonfiction authors on the ART-R in Study 3. In the present research, age positively predicted fantasy and altruism (Study 1) and positively correlated with fiction-exposure (Study 2; in line with Kidd & Castano, 2017a). These findings highlight the importance of accounting for the impact of gender and age in research examining the effects of fiction-exposure.

Verbal abilities can influence both engagement with fiction and performance on some tests of empathic abilities. Studies have shown that the development of mentalising abilities is linked to language skills (e.g., Milligan, Astington & Dack, 2007), and less skilled readers can

find it more difficult to become immersed in stories (Nell, 1988). The SST requires interpretation of a stimulus text, and the Yoni test (along with other established measures like the RMET) require the application of linguistic terms. However, the SST text is easy to comprehend (Flesch Reading Ease Score = 93 on a scale of 0-100 where higher scores indicate easier texts) and the Yoni test predominantly relies on the selection of images (requiring simpler mental state terms than the RMET). As proposed in Chapter 2, Section 2.3, researchers can account for demand characteristics through the concomitant use of both language-based and pictorial empathy tasks.

Dispositional factors such as openness to experience (Mar et al., 2009), and motivational factors such as social needs (e.g., Gabriel & Young, 2011; Mar et al., 2009), mood (McQuail, 2010; Zillman, 1988; see also Mar et al., 2011), and the need for cognition (Shim, Lim, Jung & Shin, 2018) may play a part in fiction-selection, immersion, and their effects. Exploratory results from Study 1 (Appendix B) showed that differences in motivation can impact relationships between medium and empathy. For example, watching a film to improve mood was associated with higher empathic concern, but watching a film due to feeling one "should" have seen it (perhaps due to social pressures) was negatively associated with empathic concern. Future research may identify interaction effects of presentation channel with individual differences and motivational factors. This could contribute to an understanding of the why particular formats are selected and the effects of those selections. Subsequently, this could inform interventions aiming to present stories in ways that will maximise their impact on empathy.

8.3.2 Narrative mode

"One reads alone, even in another's presence" (Calvino, 1980/1998, p. 147), although reading was not always a solo exercise. The proliferation of printed books that could be read in private, and repeatedly, led to the development of reading as an individualistic pastime

(Lagerfeld, 1986). This facilitated intimate emotional engagement with texts, and subjective interpretation of their messages. Now, again, the fiction landscape is changing. Fiction is no longer confined to the page or to transient oral tellings. Today's stories can be viewed repeatedly onscreen and interacted with using virtual headsets. Rose (2011) argued that, "people have always wanted to in some way inhabit the stories that move them. The only real variable is whether technology gives them that opportunity" (p. 88). Medium is a probable antecedent of immersion (van Laer et al., 2014; see also Mangen & Kuiken, 2014) and so immersive processes may vary across media formats. It has been suggested that the process of receiving narrative content and identifying characters has become so ingrained that the medium through which the content is presented functions merely as a transmitter (MacDorman, 2019). However, the present findings provide evidence of multiple dimensions along which both fiction media and content can have effects (see also Gentile et al., 2009, who made the same argument based on their study of video games).

Some variation in empathic abilities appears unique to processes associated with fiction. This may reflect functional differences in narrative versus exposition, or the uniquely open-minded style of engagement entailed by framing a text as fictional (Keen, 2007). Further research is required to model the impact of specific textual features within genres, though the present research indicates that features of both medium and genre that invite active engagement may particularly support empathic and prosocial development.

8.3.2.1 Participatory modes of fiction-engagement. The difference between fiction-engagement and the flight simulation metaphor (see Chapter 3, section 3.4.1; Mar & Oatley, 2008; Oatley, 1999; 2011b; Oatley & Djikic, 2017), is that the outcomes of fictional interactions are predetermined. With fiction, the reader (or viewer) cannot test out the consequences of their actions as they can in the artificial cockpit. However, trial-and-error can be facilitated in interactive modes of fiction-engagement such as video games, interactive

films, acting and improvisation (a script may be defined, but actors can test the consequences of their delivery). Participation may augment the effects of the social simulation, and this represents an exciting area for further research.[23] For example, it would be interesting to explore the extent to which participation in fictional simulations leads to behavioural outcomes compared to the more passive, traditional modes of reading or viewing. In the present research, there was some evidence of a positive relationship between acting experience and altruism. However, portraying a character can be approached using different methods that invite different degrees of imagination and emotional engagement (e.g., Gallagher & Gallagher, 2019). Future work with actors may present a unique opportunity to gain insight into the qualitative differences between ways of becoming immersed in stories and characters, and how far these may recruit and enhance mentalising and experience-sharing components of empathy.

8.3.3 Latency

The present research provided evidence for positive associations between fiction and empathic abilities but limited support for immediate causal effects. It may be that causal effects of fiction on empathic abilities occur via repeated exposure to fiction. Alternatively, fiction may influence empathic abilities, but effects may only emerge after an incubation period (e.g., Hakemulder, 2004). Drawing on the narrative persuasion literature, research has suggested that the persuasive effects of fiction may be progressively internalised over time ("sleeper effect"; Appel & Richter, 2007; Kumkale & Albarracín, 2004). In terms of effects on empathy, research has indicated that playing prosocial video games can lead to later

---

[23] Focalisation and immersion could also be manipulated in video games: A preference for inverting the *Y*-axis (which results in the control pad's "up" action moving the avatar backwards instead of forwards) has been shown to relate to immersion (Frischmann, Mauloua & Procci, 2015) and may influence identification with the avatar. More traditional modes, too (e.g., reading, film-viewing), are increasingly interactive due to the internet facilitating commentary from both laypeople and critics (see Rose, 2011). The intention to take part in online communities and express views in public forums might motivate a distinctive–perhaps analytic–style of engagement.

prosocial behaviour (Gentile et al., 2009), and Bal and Veltkamp's (2013) study of the effects of transportation into fiction presented some evidence of an increase in empathic concern following a two-week incubation period. Therefore, effects of fiction, including those influenced by immersion, may be not be observed through immediate testing (or may be so small as to require high-powered observations). The lack of evidence for long-term effects has led some authors to call for further research (e.g., Mar et al., 2011). Considering that the evidence-base for correlational effects is far more developed than for causal effects (and this contrast is visible both in the extant literature and in the present research), causal, longitudinal studies would be of great benefit to a research agenda aiming to assess the durability of causal effects.

8.3.4 Contextual variables

As shown in Study 2, differences in target mental state content, the person being empathised with and the strategy recruited, may all influence empathic accuracy. Empathy judgments can be biased: in Study 2, perceivers tended to be conservative in their estimates of targets' emotional ratings, and the directions in which biases occurred (based on the positive or negative valence of estimates) varied depending on empathic inferencing strategy condition. Observed biases may differ depending on target characteristics such as gender (Garrido & Prada, 2017), expressiveness (Allport & Vernon, 1933), and similarity to the perceiver (Adams et al., 2010; Matsumoto et al., 2009). They may also differ depending on the type (cognitive or affective) and complexity (e.g., first- or second-order) of mental state content, and mode of presentation (e.g., naturalistic versus pictorial; Richter et al., 2010). Examining the influence of context variables on the deployment of inferencing strategies could facilitate a nuanced approach to indexing individuals' empathy. Similarly, the prosocial behaviour component was measured in Study 1 using a self-report measure of altruism. This scale contains questions about helpful acts towards strangers, acquaintances and charities.

These scenarios can differ in terms of social visibility; for example, a charity donation could be made privately on the internet, whereas giving money to a stranger, or donating blood, may be more public. Future research could address the differences between prosocial acts and how far they may be influenced by fiction engagement.

8.4 Methodological Limitations and Recommendations for Further Research

There were several limitations in the present research project; some of which have been identified in the study-specific discussions above. This section presents more general methodological limitations and recommendations for addressing these in future research.

8.4.1 Limitations of measures

8.4.1.1 Measures of empathic abilities. The self-report empathy scales (used as criterion variables in Study 1 and control variables in Studies 2 and 3), and immersion scales (Study 3) required participants to accurately reflect on their own thoughts and behaviours and so they may be subject to bias. This can take the form of socially desirable responding, where participants provide answers that adhere to positive social norms (e.g., presenting oneself as altruistic), or from a disproportionate sense of one's own skills (Kidd et al., 2016, noted that the predictability of popular fiction characters can engender a sense of competence in empathy, and so this may be particularly relevant to the genre correlations). Alternatively, participants may try to support the researcher's aims, they may generally provide positive (acquiescing) or negative (dissenting) responses, or they may simply not pay attention to the questions. In order to remedy these limitations, irrelevant items were included within self-report scales to mask the hypotheses, attention check questions and reverse-scored items were included, and internal reliability was assessed using Cronbach's alpha, which enables the detection of inconsistencies within response sets.

The empathic accuracy task created for Study 2 relies on the experiencers' abilities to accurately reflect on, and consistently index, the valence and intensity of their emotional

responses to pictures. This approach presents greater ecological validity than, for example, schematic faces such as the Yoni test used in Study 3 (e.g., Dobs, Bülthoff & Schulz, 2018). However, as participants did not directly estimate the experiencers' affective responses, rather they estimated the ratings provided by the experiencers, this comes at the cost of an objectively "correct" set of values against which to verify participant estimates. Furthermore, the extent that mentalising and experience-sharing processes were induced by the empathic accuracy task (Study 2) is not certain. As participants in the experience-sharing ("simulation") condition had access to photographs of the experiencers during each trial, they may have based their estimates on their knowledge of other similar people. In sum, they may have employed a mentalising approach. Conversely, participants that viewed the videos may have drawn on their own responses to the experiencer reactions. Nonetheless, the association between accuracy and lifetime fiction-exposure, as well as the observed difference in directions of error, merit further investigation. Future research could remedy these issues by assessing concurrent validity with established mentalising and experience-sharing measures.

8.4.1.2 Measures of fiction. ARTs establish familiarity with specific selections of authors as indirect measures of exposure to fiction, but they do not reflect depth of engagement or enjoyment. A one-point increase for recognising the name "José Saramago" is attained whether participants have read and re-read Saramago's full catalogue, skimmed through one of his books, or simply heard his name in conversation. External factors such as age, location, cultural differences, education levels and experiences (e.g., differences in school syllabuses) are likely to affect recognition rates. Nevertheless, previous research has lent support to the validity of ARTs, showing that familiarity with authors correlates with measures of reading ability and habits (Mol & Bus, 2011; Rain & Mar, 2014; Stanovich & Cunningham, 1992; Stanovich & West, 1989; Stanovich et al., 1995; West et al., 1993). The validity of the FMET, on the other hand, would need to be established through further

research. However, the observed associations between media-exposure and respective media preferences (Study 1) indicate its promising capacity to characterise fiction habits.

The ARTs and FMET used in the present studies were not analysed at the level of their genre subcomponents. Future research could index the FMET along genre dimensions in order to ensure that it contains sufficient breadth. This approach could facilitate an examination of the effects of genre-exposure across different media (for examples of ARTs used to measure genre-exposure, see Black et al., 2018; Kidd & Castano, 2017a). Furthermore, the FMET was restricted to fiction, and as fiction and nonfiction print-exposure tend to correlate (Mar et al., 2006, 2009), nonfiction dimensions could be usefully developed. Studies 2 and 3 focused on the effects of print, and Study 1 examined print, film and play-exposure. While the FMET represents a development of the ART, it still provides only a narrow view of the variety of modes though which fictional stories can be accessed. Study 1 indicated that media were differentially associated with empathic abilities in terms of both strength and direction (e.g., positive associations with preference for novels and negative associations with preference for TV). Future research could expand the FMET in order to examine these differences in more detail and to better reflect the multiplicity of the fiction landscape.

Study 3 failed to vary self-report immersion using a cognitive load manipulation and, compared to previous studies (e.g., Green & Brock, 2000), immersion scores were generally low. Features specific to the selected text may have attenuated effects on immersion dimensions (e.g., the specific characters and themes depicted). Employing a variety of texts would allow future researchers to examine effects and generalise beyond a single passage (this introduces heterogeneity when comparing genres between groups, but nested statistical models can circumvent this issue; Panero et al., 2016).

8.4.1.3 Statistical power and reliability. Multiple comparisons were reported in the present research (the correlation matrices in Studies 1 and 3, and tests conducted on both raw and absolute scores reported in Study 2 were not corrected for multiplicity). However, correcting for the reporting of both raw and absolute scores in study 2 (by applying a Bonferroni correction) did not affect the main pattern of results, and patterns of association tended to internally replicate (for example, the pattern of correlations in Study 3 Experiments 1 and 5). Other statistical tests incorporated corrections for multiple comparisons (e.g., ANOVA and Bonferroni-corrected post-hoc comparisons). Violations of test data assumptions were reported and addressed.

The causal effects of immersion in Study 3 were likely underpowered, and future research could explore the impact of the cognitive load manipulation using a larger sample. The selected fictional text was a short excerpt from a longer novel. Research has indicated that sensorimotor simulation is higher for full narratives than unconnected sentences extracted from narratives (Kurby & Zacks, 2013) and, similarly, immersion may be weaker for short extracts compared to full narrative arcs (containing antecedents, consequences and resolutions to events). Using short passages precludes generalising the results to naturalistic reading, but future longitudinal research employing full narratives may very usefully address this issue. This would provide a more ecologically valid perspective on fiction leisure-reading and contribute to clearer estimates of the sizes of fiction-reading effects.

In extant studies using short passages, researchers have tended not to alter published original works of fiction (as was the case with the Study 3 stimuli), an approach that offers ecological validity but comes at the cost of experimental control (Willems & Jacobs, 2016). Some authors have incorporated original and altered versions of texts (e.g., Hakemulder, 2004), which can facilitate the isolation of specific textual elements and measurement of their effects (e.g., focalisation and foregrounding devices such as imagery and metaphor). This

approach may be too resource-heavy to use in longitudinal designs, but cross-sectional research employing textual manipulations, combined with longitudinal approaches examining naturalistic reading, may work contemporaneously to clarify the strength of the impact of fiction-reading.

8.5 Making Sense of the Results: How Does Fiction-engagement Benefit Empathy?

> Humans had built a world inside the world, which reflected it in pretty much the same way as a drop of water reflected the landscape. And yet ... and yet ... Inside this little world they had taken pains to put all the things you might think they would want to escape from——hatred, fear, tyranny, and so forth. Death was intrigued. They thought they wanted to be taken out of themselves, and every art humans dreamt up took them further *in*. (Pratchett, 1988, p. 328)

Pratchett observed that people use fiction to explore challenging themes. This invokes Keen's (2007) suggestion that fictional worlds function as safe zones wherein readers can explore themes and scenarios without real-world consequences. Nettle (2006) proposed that human relationships are central to fiction, and Mar and Oatley (Mar & Oatley, 2008; Oatley, 1999, 2016) suggested that the function of fiction is to support social development by simulating social experiences. This theory offers a solution to the paradox that fictional characters must sufficiently resemble real-world others in order to generate emotion while, at the same time, the reader knows that they are not real (Radford & Weston, 1975). Fiction simulations are not "direct impressions of the world" (Oatley, 2011b, p. 16), rather they are dreamlike imitations of real events (Oatley, 2011b). Veering off the runway in a flight simulator would induce a similar, but not identical, emotional reaction to doing so in a real plane. Similarly, the scenarios explored in fiction are inconsequential, and may invite emotional engagement for precisely this reason (Keen, 2007). Just as flight simulators function to develop pilots' competencies in a cockpit, fictional narratives may develop readers' competencies in matters of social life.

The current research examined three overarching questions: (i) what is the nature of relationships between different narrative modes and empathic abilities? (ii) what is the nature of relationships between fiction and mentalising and experience-sharing strategies for empathic accuracy? (iii) what is the role of immersive processes in fiction effects on empathy? These questions were addressed in the current studies with varying levels of success.

8.5.1 What is the nature of relationships between different narrative modes and empathic abilities?

Addressing the first question, the collective findings indicated that both media and genre differences are associated with empathic abilities including altruistic behaviour, which may represent a downstream consequence of fiction effects on empathy. Print-exposure (reading) and engaging with comedy or postmodern genres may particularly support empathic skills, due to the active, writerly modes of engagement that they entail. However, more formulaic narratives, too, may benefit empathic abilities, potentially via different routes. While most of the extant research evidence indicates that stories can impact empathy through the development of social processes (Mar, 2018a), the findings that different thematic genres relate to different empathic faculties suggests that people may also learn and apply knowledge accumulated through engaging with specific social themes and scenarios.

8.5.2 What is the nature of relationships between fiction and mentalising and experience-sharing strategies for empathic accuracy?

Addressing the second question, fiction-exposure was found to be associated with empathic accuracy both when sharing in the same experience as a target and when interpreting a target's facial expressions. Fiction-exposure was not found to uniquely support one mode of inferencing over the other, rather it may support both routes to empathic

accuracy. It remains to be seen how far fiction-engagement can causally enhance each strategy.

8.5.3 What is the role of immersive processes in fiction effects on empathy?

Addressing the third question, the impact of immersive processes in empathic accuracy remains unclear. Immersion dimensions appear to relate to self-reported empathic traits (mentalising, experience-sharing and concern dimensions), as well as to behavioural measures of empathic accuracy (for cognitive and affective content), and so these relationships merit further enquiry. This would help to determine how far imaginative and affective dimensions invoke social processes, support the understanding and retention of social content, or facilitate both non-exclusively (see Mar, 2018a). This, in turn, could contribute to debate on the extent that real-world empathy requires the self and other to remain distinct and involves imaginative versus isomorphic processes (e.g., Cuff et al., 2016; see Gallagher & Gallagher, 2019, for an overview).

8.5.4 The question of causation

In line with the literature reviewed in Chapter 3, the present research provided evidence that fiction-exposure and empathic abilities are positively associated, while the causal evidence was far weaker. It could simply be the case that people who are higher in empathic abilities—or particular empathic components—tend to engage more with fiction. If so, fiction may simply represent a "virus of the mind" (Brodie, 2009) that has hijacked the human capacity for multiperspectivity and ability to decipher complex social scenes but offers nothing by way of improvement. Pinker captured this idea in his discussion of "pleasure technologies":

> We enjoy strawberry cheesecake, but not because we evolved a taste for it. We evolved circuits that gave us trickles of enjoyment from the sweet taste of ripe fruit, the creamy mouth feel of fats and oils from nuts and meat, and the coolness of fresh water. Cheesecake packs a sensual wallop unlike anything in the natural world because it is a brew of

megadoses of agreeable stimuli which we concocted for the express purpose of pressing

our pleasure buttons. Pornography is another pleasure technology. . . . the arts are a third.

(Pinker, 1997, p. 525).

The view that fiction can be beneficial—cognitively and socially—is seductive. It

means that engaging with stories for leisure is not merely entertaining or enjoyable; it is

healthy too (Nell, 1988, coined the term "ludic readers" to refer to people that read for

pleasure and relaxation more than once a week, p. 7). Of course, different fictions are not

always pleasurable but are compelling (e.g., the horror novel *Dracula,* Stoker, 1897/2019; or

the melodramatic film *Dancer in the Dark,* Von Trier, Jensen & Windeløv, 2000). However,

extending Pinker's metaphor from the arts, broadly, to fiction, specifically, risks being overly

reductionist. It implies that longstanding literary classics have no greater intrinsic value than

daily soap operas. The idea that fiction-engagement is merely a symptom of people's

cognitive capacities does not explain the different outcomes observed in causal studies that

have contrasted text genres (e.g., Kidd & Castano, 2013, 2018b; Kidd et al., 2016; Van Kujik

et al., 2018); it may tell part of the story, but it does not provide the complete picture.

Part of the picture concerns the beliefs of fiction–empathy researchers themselves. In

literary theory, empathy has always represented an assumed, typical reaction to a literary text,

helping to distinguish literature from other text genres (Keen, 2007; Pinotti & Salgaro, 2019).

The idea that fiction can function to support real-world empathy is perhaps particularly

appealing to researchers who, identifying as writers or readers themselves, are drawn to study

the topic, and so the extant literature may have been biased by studies generated by

"bibliophiles" (see Koopman, 2018). As author Zadie Smith recently commented, "if fiction

had a belief about itself, it was that fiction had empathy in its DNA, and that was the product

of compassion […] This was what fiction believed about itself, but like all beliefs not a little

of it was always wishful thinking" (Smith, 2019, p. 9).

This (perhaps "wishful") thinking not only represents what fiction might believe about itself, but also what psychologists might believe about fiction. Several empathy researchers have employed narrative-based stimuli as naturalistic alternatives to pictorial and vignette tasks (see Chapter 2 for examples). Implicit in this approach is the assumption that fiction-induced empathic processes are equivalent to real-world empathic processes. Psychological research in the field of fiction's effects on empathy is also based on this supposition, which Pinotti and Salgaro (2019) argued is a misinterpretation perpetuated by psychological researchers such as Mar and Oatley. However, the theory that fiction supports empathy—even that it functions to do so—does not necessarily assume that fictional empathic processes and real-world empathic processes are the same, rather it assumes that they are related. The simulation is considered to imitate the real world sufficiently well that some of the same cognitive mechanisms are activated, and that feelings experienced during reading reflect, but are not the same as, feelings experienced during similar real-world events (e.g., Oatley, 2011b). Neuroscience research outlined in Chapter 3 has lent some support to this theory (also see Mar's, 2011, meta-analysis). The extreme argument that fiction-induced empathy has no real-life consequences falls short of explaining the robust associations between fiction-exposure and empathic abilities shown in the scientific literature. The degree to which the "peculiar" aspects of "fictional empathy" (Pinotti & Salgaro, 2019, p. 150)—empathic experiences unique to literary-engagement—lead to real-world psychological or social change, however, remains a central question for future research to continue to explore.[24]

While the idea that fiction has no effect on empathy may be disappointing to "bibliophiles" (Koopman, 2018), the idea that fictional empathy is not unique or mystical, but is underpinned by the same processes as run-of-the-mill, everyday empathy could be just as

---

[24] This question may differ across modes of fiction. For example, it is not yet known whether audiences process actors in performances like they process fictional characters in novels, whether they perceive them as essentialist representations of traits and states, as real people telling lies, or as complex everyday people (Goldstein & Filipe, 2018).

unappetising. Robust causal evidence of the behavioural consequences of fiction-engagement is required to substantiate the claim that fiction is functional, and that its central function is to potentiate empathic growth. The present literature appears to indicate a small, positive effect overall, but is sufficiently mixed as to warrant further investigation.

8.5.5 Fiction and empathy and the assumption of something "good"

For every *Uncle Tom's Cabin* there is *Birth of a Nation.* For every *Bleak House* there is *Atlas Shrugged* … Every single one of these fictions plays on its readers' empathy: not just high-minded writers like Dickens, who invite us to sympathize with Little Dorritt, but also writers of Westerns, who present poor helpless colonizers attacked by awful violent Native Americans (Landy, n.d., as cited in Bloom, 2016, pp. 48-49).

Research examining how fiction may causally enhance empathy is predicated on two assumptions. The first is that fiction is "good" and may have the power to change things for the better. The second is that empathy is "good" and that increasing it is inherently beneficial. Keen (2007) pointed out that if the premise that fiction can alter readers' characters for the better is accepted, so must be its opposite: that fiction can also "possess darker powers" (p. 25). If fiction can change people's attitudes and beliefs (e.g., Green & Brock, 2000), it may also have the power to propagate stereotypes, alienate, distance and dehumanise (research examining fiction's antisocial outcomes has predominantly focused on video games, e.g., Fraser, Padilla-Walker, Coyne, Nelson & Stockdale, 2012; for a recent meta-analysis of effects on aggression, see Prescott, Sargent & Hull, 2018; for an overview, see Rose, 2011). Fiction-engagement could have a negative social impact by providing a means to escape pressing social issues; bystanders can perform a "mental disappearing act" through reading in order to avoid taking prosocial action (Keen, 2007, p. 25). To prevent this, some modes of fiction, such as Epic Theatre, purposefully employ techniques designed to emotionally

distance audiences in order to provoke rational and critical evaluation of the ideological issues presented.[25]

Real-world empathy does not automatically, directly or inevitably lead to helping (e.g., Eisenberg & Strayer, 1987), and the behavioural consequences of fiction-induced empathy are equally uncertain (e.g., Keen, 2007). While the fiction simulation may function as a mental, "moral laboratory" (Hakemulder, 2000), if fiction is used as an escape, this could come at the cost of real-world prosocial action. Study 1 showed that self-report altruism was positively predicted by fiction-exposure (and previous research has provided some evidence of causal effects, e.g., Johnson, 2012; Johnson et al., 2013; Koopman, 2015). Further research examining causal effects on behavioural measures is required to substantiate the association between fiction, empathy, and prosocial, ethical behaviours.

The concept of empathy has become ideologically entangled with ethics. In response to the "empathy hype" (a wave of interest in empathy in aesthetics and the humanities over the past 20 years; Pinotti & Salgaro, 2019, p, 144), Bloom (2016) and others (e.g., Breithaupt, 2019) have argued that empathy can be detrimental. Echoing Pinker's (1997) depiction of the arts as providing the "sensual wallop" (p. 525) of cheesecake, Bloom (2016) designated empathy as, "sugary soda, tempting and delicious and bad for us" (p. 13). In other words, what feels good in the moment, might not be so good long-term. Bloom suggested that empathy can be misguided, short-sighted and biased, and argued instead for a utilitarian approach to "rational compassion" (his prior research having shown that compassion dissociates from experience-sharing; Bloom, 2017). He presented an example in which

---

[25] Bertolt Brecht (1898-1956) wrote several plays that functioned as forums for political and social issues to be scrutinised, camouflaged as entertainment. For example, *The Resistible Rise of Arturo Ui* (Brecht, 1941/1993) chronicled the rise of a Chicago mobster as a satirical allegory for the rise of Hitler and the Nazi party prior to World War Two. Emotional engagement was discouraged through techniques such as breaking the "fourth wall" (the imagined wall separating actors from spectators, through the characters directly addressing the audience, for example), as it may have reduced spectators' abilities to elaboratively reflect on the political commentary of the piece.

someone high in empathy might seek a "warm glow" (p. 99) by donating small amounts of money to multiple charities. He pointed out that this approach could be financially detrimental to the charities because of the associated processing costs and resources spent following up on those donations. He cautioned that, "if you want to harm some organization that supports a cause you object to, one mischievous way to do so is to send them a $5 donation" (p. 99). However, Bloom's scenario does not present a problem with empathy *per se*, rather it presents a problem with defective knowledge about an action's outcome. Empathy is not really pitted against reason, because empathy is not given the chance to do its job—empathy does not have all the information. Rational compassion, too, may find itself lacking, in the absence of information with which to make a rational choice.

There are two broader issues with the idea that reason represents an antidote to empathy. The first issue is the implication that one prosocial act is objectively more valuable than another (which serves to further conflate empathy and ethics) in relation to complex and contextual social issues. This is illustrated in Bloom's example of the "considerable debate" (p. 99) that surrounds Western aid provisions to developing nations: on the one hand this can be considered a helpful and socially conscious act; on the other, it could have negative, downstream consequences for local economies. The second issue, of practical importance to the current research agenda, is that Bloom's (2016) argument defines empathy in terms of its isomorphic, affective component (see Zaki, 2017). This version of empathy may be characterised by feelings of distress, which activate avoidance rather than helping. As described in Chapter 3, distress tends to be self-oriented and concerned with the alleviation of one's own suffering, rather than with that of another, which distinguishes it from the other empathy dimensions (e.g., Baron-Cohen & Wheelwright, 2004; Davis et al., 1999; Penner et al., 1995). Whereas Bloom (2016) argued that affective empathy creates a "spotlight" (p. 9) which biases the direction of moral efforts, Zaki (2017) suggested that the spotlight can be

controlled and usefully redirected. Emotion regulation abilities, for example, can direct empathisers toward non-isomorphic empathic concern and away from personal distress (McCall & Singer, 2013). Zaki conceded that the empathy spotlight can be misdirected, biased, and can foster both moral and immoral outcomes. He called for scientists to consider when empathy most aids empathisers and targets, in order to "help people tune their emotional lives accordingly" (Zaki, 2017, p. 60).

Fiction-engagement represents one method for tuning emotional life, though it may not always be a force for good. Some content, and some modes of engagement, can have negative social outcomes. For example, fictional empathy may perpetuate biases through over- or under-representation of individuals and groups (e.g., Iguarta, Barrios & Ortega, 2012; see also Smith, 2019), which may bolster real-world empathy biases. However, if fiction can foster the empathic components that engender prosociality (e.g., concern rather than distress), it may support the ability to flexibly redirect the empathy "spotlight" in order to promote helpfulness, cooperation and equity.

Future research may usefully investigate the processes through which fiction can impact real-world helping. It may be that accumulating social information through fiction generates an awareness of the experiences of particular people or groups, which can alter attitudes and, consequently, behaviours towards those individuals and groups (a route that may overlap with persuasion). Alternatively, the processes of engaging with fiction may enhance empathic abilities in general, leading to an increased tendency toward prosocial behaviour. These alternative routes could be examined using stories that vary in terms of prosocial or antisocial characters and outcomes. For example, the story composed for Study 3 presents a positive outcome to an altruistic act performed by the protagonist. A second version of the story could depict a negative outcome to the same act. If fiction alters behaviour through the accumulation of knowledge informed by the content of the story, participants would be

expected to exhibit different levels of altruistic behaviour depending on the version of the story that they had read. In contrast, if story-induced empathic processes lead to altruism, behaviours should not differ depending on the outcome of the story. Such questions highlight the importance of the multidimensional approach to fiction-induced empathy in clarifying its consequences, costs and benefits.

8.6 Contribution and Implications

8.6.1 The multidimensional approach

The measures employed in studies of empathy "tell us not about the nature of empathy but about certain ways in which empathy can be achieved" (Smith, 2017, p. 709). To adapt McLuhan's (1964/1994) statement (the "medium is the message"), in psychological research on empathy, for the most part, the *measure* has been the message. The multidimensional framework adopted in the present research enabled a reconsideration of previous findings of fiction effects on empathy in terms of the modes of measurement employed, and the empathic processes probed by those measures. This approach enabled results from existing correlational and causal studies to be categorised according to mentalising and experience-sharing, concern and prosocial behaviour components. In so doing, it offered some explanation for the currently mixed findings in the field of fiction effects on empathy—heterogeneity across fiction stimuli and empathy measures—and promoted a granular approach to modelling their relationships.

The research lent supporting evidence to the hypothesis that fiction-exposure and empathic abilities are positively associated. Fiction-exposure was shown to positively relate to mental state content, measured using both self-report and behavioural tools. This result aligned with previous findings that nonfiction, too, positively relates to empathic abilities, but that the effect is not as large as for fiction, indicating that there is something special about fiction in terms of its impact on empathic processes. The research extended this line of

enquiry by showing positive relationships between empathic abilities, print and theatre exposure, as well as negative relationships with preferences for TV and film. Genre preferences were differentially associated with empathic abilities. Due to the exploratory nature of Study 1, the results for media and genre effects were preliminary. However, the FMET tool laid the groundwork for further research employing pluralistic fiction-exposure measures in order to explore the effects of a wider range of media and genres.

The approach to measuring empathic accuracy taken in Study 2 enabled the relative efficacy of mentalising and experience-sharing routes to be examined. The nuanced approach to measurement, employing both absolute scores to assess magnitude, and raw scores to assess directions of effects, extended Zhou et al.'s (2017) prior research. It facilitated the discovery that mentalising and experience-sharing strategies lead to similar levels of error, but in opposite directions. This difference did not moderate the relationship between fiction-exposure and empathic accuracy, suggesting that fiction-exposure may strengthen empathic accuracy via both routes. Future research may establish whether this finding replicates. This represents the first study to directly compare the relationship between fiction-exposure and mentalising versus experience-sharing routes to empathic accuracy. The systematic approach taken in Study 3 contributed limited evidence of a positive effect of a reading task on empathic accuracy and some evidence of relationships between empathy and immersive processes. A method for manipulating immersion in a popular narrative was identified, which may be tested in future research using different story texts. Study 3 highlighted the multifaceted nature of immersion, which includes imaginative processes, identification, and emotional engagement, and which may differentially relate to empathy components.

8.6.2 Influential variables

The research highlighted several variables that can impact empathic accuracy and the relationships between fiction and empathic abilities. These include both demographic

variables such as gender and age, contextual variables, such as characteristics of the target person, valence of the mental state content and empathic strategy employed. The latter, in particular, was shown to influence not the magnitude, but the direction of bias generated when estimating another's emotional experience. Further research may fruitfully develop this line of enquiry, in order to further clarify the efficacy of mentalising versus experience sharing inferencing strategies and the role of fiction-exposure in their development.

8.6.3 Measures developed

Taking a multidimensional approach necessitated the employment of a range of measures, and three tasks were designed or developed as part of this research: the FMET was created for Study 1, which facilitated a more comprehensive assessment of fiction-exposure than previous ARTs. An empathic accuracy task was produced based on Zhou et al. (2017) for Study 2, which extended the research agenda by examining the impact of inferencing strategy on accuracy for the emotional experiences of other participants. The ART-R (Mar et al., 2006) included recently published authors (for Studies 2 and 3). These materials are available for use in future research. Additionally, three texts, including one written for the study (Appendix F), were subjected to analyses of readability and immersiveness across three dimensions. These texts and data may be useful for future researchers selecting and working with narrative stimuli.

8.6.4 Implications for knowledge of empathy and fiction

In Chapter 2 it was suggested that individual differences in fiction-engagement processes should be taken into consideration by researchers using narrative-based empathy tasks. The current investigation of the influence of immersion dimensions on empathic accuracy task performance was inconclusive. However, the research did reveal differences in levels of exposure to, and outcomes of, engaging with fictional film and print media. This suggests that individuals may respond differently to empathy tasks presented using film or

print (e.g., the MASC or SST; see Chapter 2), and so furthering the exploration of fiction effects across media channels would not only inform research in fiction, but it also has implications for the use of fiction stimuli in empathy research more broadly.

Chapter 3 invoked De Vignemont and Singer's (2006) observation that there are "nearly as many definitions of empathy as people working on the topic" (p. 435). Cuff et al. (2016) argued that determining an agreed, single definition of empathy would support researchers to develop measures and compare results across studies. Empathy is conceived in the present research as multidimensional phenomenon: it cannot be described in terms of a singular psychological state or process; rather a range of processes may be recruited in service of empathic accuracy (in line with Smith, 2017). Empathic accuracy represents an epistemological phenomenon: an emergent form of knowing that is constituted through the recruitment and interactions of prior knowledge and present processes. Thus, it has been argued in the current research that a unidimensional definition of empathy is not achievable due to the complex nature of the phenomenon and the range of constituent lower and higher order cognitive and affective processes. Moving forward, the field may begin to converge on a multidimensional definition that encompasses the range of psychological processes involved in understanding others.

Fictional narratives recruit and potentially strengthen processes associated with empathy, and it is for this reason that they have been widely used as naturalistic stimuli in measures of the ability to accurately attribute mental states. While it is generally accepted that people mentally simulate events in fictional narratives, the nature of the fiction simulation is underdetermined (Willems & Jacobs, 2016): it remains unclear which aspects of fictional worlds people simulate, and the cognitive mechanisms through which these forms of simulation may influence empathy. Taking a multidimensional approach to fiction, as well as to empathy, calls for an in-depth study of specific narrative features and forms that entail the

processes associated with empathy dimensions. Due to the complexity of the question of how fiction may benefit empathy, progress will be incremental rather than immediate. Continued research will gradually contribute evidence that converges on the antecedents, processes and consequences of fiction-developed empathy. In turn, this will shed light on empathic inferencing processes, described by Ickes (1997) as the "second greatest achievement" (p. 2) of the human mind after consciousness. The fiction landscape is multidimensional and changing, and the present research represents a step towards a more nuanced and comprehensive understanding of its social impact.

8.6.5 Social implications

The current research has implications for supporting empathic development in NT adults and groups with characteristic deficits. In a study of American college students, self-reported empathic concern and perspective-taking were shown to have dropped in recent decades (1979-2009; Konrath, O'Brien & Hsing, 2011). Zaki (2019) proposed that this may be due, in part, to an increase in the number of people living alone, and the anonymous, urban communities in which they live. The results of the current study, as well as results from previous research, indicate that engaging with stories can develop the ability to accurately interpret the thoughts and feelings of others, feel concern for others and feel motivated to alleviate their suffering. In this way, fiction-engagement may buffer against a decline in empathy. These findings could be particularly meaningful to individuals and organisations working in person-centred contexts such as social care which, by definition, require empathic skills. Taken forward, the multidimensional approach may also inform interventions with groups that experience social difficulties by systematically treating the specific deficit (such as emotion identification in alexithymia, mentalising in autism and affect-sharing in psychopathy; e.g., Freedman et al., 2013). Present and future evidence of the psychological impact of stories may benefit organisations proposing—and seeking funding for—arts-based

interventions aiming to promote prosocial development and positive interpersonal relationships.

8.7 Moving Forward

Fiction simulations may increase knowledge of the experiences of distant others and strengthen the processes that people use to understand them (e.g., Mar et al., 2018a). Both of these effects may serve to reduce the misconception that others' experiences are very different to one's own, and to expand people's empathy—or their tendency to deploy empathic strategies—beyond the immediate ingroup (see Singer, 1981/2011). There is some evidence that readers internalise the social identities that they discover in fiction. Gabriel and Young (2011) found that readers tended to identify more with the social group that they had read about, even when groups were fantastical (in this case, vampires and wizards). The mechanisms engaged through fiction may increase flexibility around self-other distinctions and this flexibility may extend to other aspects of cognition. For example, experience with fantasy and science-fiction genres, which require consumers to imagine new environments governed by unusual physical laws and social norms, has been linked to increased flexibility in both moral and physical judgments (Black et al., 2018).

The extent to which empathy for the ingroup—the "empathy circle" (Pinker, 2011; Singer, 1981/2011)—may be expanded through engagement with fiction is yet to be fully explored. The introduction to this thesis made reference to several organisations using arts engagement practices designed to develop social skills and to improve social experiences. These organisations have focused on the development of empathy and prosocial behaviours in different human groups (e.g., children, prisoners, medical and social care workers). However, future research may identify the potential effects of fiction-induced empathy for other animals ("interspecies empathy"; Pinotti & Salgaro, 2019, p. 151; see also Grandin & Johnson, 2009) and the environment. For example, researchers in sustainability (e.g., Chapman, 2015) have

begun to foreground empathy as an antidote to planned obsolescence—a feature of design that limits the life of a product and leads to waste—and arts organisations have developed programmes aimed at engendering positive behaviours toward animals and nature (e.g., Marks, Chandler & Baldwin, 2016). The recent proliferation of environmental and eco-fiction genres may be shown to influence attitudes and behaviours towards pressing environmental issues (e.g., David, 2016), through foregrounding the plight of animals and natural environments. Empathy targets in fiction need not be human to be *humanlike* (see Mar, 2009; Mar & Oatley, 2008). It was established in Chapter 2 that people show an automatic tendency to attribute mental states to objects and shapes (exemplified in Heider and Simmel's, 1944, classic study). Titchener both highlighted people's affection for books and encapsulated this tendency when he posed the following question:

> What lover of books has not shifted the place of certain volumes on the shelf, because he could not bear to put good and bad, sound and trivial, side by side,—as if the *books* would feel the incongruity? (Titchener, 1909; as cited in Lanzoni, 2018; pp. 66-67)

As fictional depictions can imbue objects with humanlike characteristics, future research may establish how far this can be exploited and translated into real-world action. Exploring this in more depth would both provide information about the parameters of fiction-induced empathy processes and help to determine the potential impact of fiction-induced change at both micro individual and macro societal levels.

8.7.1 Questions for future research

Based on previous research and the current findings, three directions for future research are proposed:

1. **Model the relationships between fiction and prosocial behaviour:** Which fiction-engagement processes mediate effects on prosocial behaviour? What is the impact of contextual variables (e.g., the social context of the prosocial act)?

2. **Examine the effects of specific social content variables:** What is the impact of varying narrator (reliable, unreliable), protagonist (prosocial, antisocial) and multiperspectivity (tracking small and large character networks)?

3. **Establish the parameters of fiction-induced empathy:** What are the short and long-term effects of self-directed fiction-engagement on empathy? To what extent can fiction affect empathy and behaviours towards non-human others, objects or environments?

8.8 Summary and Conclusions

This chapter outlined the results of the empirical work conducted and discussed the findings in relation to extant research and the multidimensional approach proposed in Chapters 2 and 3. Factors affecting relationships between fiction and empathy were discussed. Limitations of the research were identified and the idea that fiction benefits empathy was critiqued in terms of the limited causal evidence and the implicit assumptions that fiction and empathy are objectively "good". Epistemic and social implications were identified and potential avenues for further research proposed.

In line with previous research in the field of fiction effects on empathy, the results indicated that fiction and empathic abilities are associated. The precise nature of the mechanisms involved remain uncertain, although the research identified some of the conditions under which fiction effects may occur. The multidimensional approach demonstrated that medium and genre are important factors in relationships between fictional content and empathic abilities. The research also showed that contextual variables, such as the identity of the experiencer, the nature of the target content, the type of stimuli employed, and the empathic inferencing strategy deployed for the task, can influence empathic accuracy and produce bias.

The effects of fiction on empathy represents a complex area of enquiry. This complexity has been reflected in inconsistencies in the ways that empathy is operationalised, as well as heterogeneity across fiction effects. However, patterns of evidence have begun to emerge

indicating that fiction-exposure is associated with the mentalising, experience-sharing and concern components of empathy, as well as with prosocial behaviour, and that a range of contextual factors can influence those relationships. These findings have implications for understanding the processes and parameters of empathy, as well as the role that fiction plays in social life. It is recommended that future research take a multidimensional approach to studying the effects of specific features of fiction in order to examine pathways to empathic components, and to elucidate the antecedents and consequences of fiction-engagement. In so doing, the benefits of fiction for empathy, and its social impact, will begin to crystallise.

References

Abel, F., Happé, F., & Frith, U. (2000). Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development. *Cognitive Development, 15,* 1-16. https://doi.org/10.1016/S0885-2014(00)00014-9

Acheson, D. J., Wells, J. B., & MacDonald, M. C. (2008). New and updated tests of print-exposure and reading abilities in college students. *Behavior Research Methods, 40*, 278-289. http://doi.org/10.3758/BRM.40.1.278

Achim, A. M., Guitton, M., Jackson, P. L., Boutin, A., & Monetta, L. (2013). On what ground do we mentalize? Characteristics of current tasks and sources of information that contribute to mentalizing judgments. *Psychological Assessment, 25*, 117-126. http://doi.org/10.1037/a0029137

Adams Jr, R. B., Rule, N. O., Franklin Jr, R. G., Wang, E., Stevenson, M. T., Yoshikawa, S., Nomura, M., Sato, W., Kveragam K., & Ambady, N. (2010). Cross-cultural reading the mind in the eyes: An fMRI investigation. *Journal of Cognitive Neuroscience, 22*, 97-108. http://doi.org/10.1162/jocn.2009.21187

Adolphs, R., Baron-Cohen, S., & Tranel, D. (2002). Impaired recognition of social emotions following amygdala damage. *Journal of Cognitive Neuroscience, 14*, 1264-1274. http://doi.org/10.1162/089892902760807258

Allport, G. W., & Vernon, P. E. (1933). The problem of consistency in expressive movement. In G. W. Allport & P. E. Vernon (Eds.), *Studies in expressive movement* (pp. 3-35). New York, NY: MacMillan.

Altman, D. G. (1999). *Practical statistics for medical research.* New York, NY: Chapman & Hall/CRC.

Appel, M., & Richter, T. (2007). Persuasive effects of fictional narratives increase over time. *Media Psychology, 10,* 113-134. https://www.tandfonline.com/doi/abs/10.1080/15213260701301194

Apperly, I. (2011). *Mindreaders: The cognitive basis of "theory of mind".* Hove, England: Psychology Press.

Apperly, I. A., Samson, D., & Humphreys, G. W. (2009). Studies of adults can inform accounts of theory of mind development. *Developmental Psychology, 45*, 190-201. http://doi.org/10.1037/a0014098

Anderson, C. A., & Dill, K. E. (2000). Video games and aggressive thoughts, feelings, and behavior in the laboratory and in life. *Journal of Personality and Social Psychology*, *78*, 772-790. https://doi.org/10.1037/0022-3514.78.4.772

Astington, J. W., & Dack, L. A. (2008). Theory of mind. In M. M. Haith, & J. B. Benson (Eds.). *Encyclopedia of infant and early childhood development, 3* (pp. 343-356). San Diego, CA: Academic Press.

Awasthi, A., & Mandal, M. K. (2015). Facial expressions of emotions: Research perspectives. In M. K. Mandal, & A. Awasthi (Eds.), *Understanding facial expressions in communication*, [Springer: EBL version], pp. 1-18. http://doi.org/10.1007/978-81-322-1934-7

Aykan, S., & Nalçacı, E. (2018). Assessing theory of mind by humor: The Humor Comprehension and Appreciation Test (ToM-HCAT). *Frontiers in Psychology, 9*, 1470. http://doi.org/10.3389/fpsyg.2018.01470

Back, E., & Apperly, I. A. (2010). Two sources of evidence on the non-automaticity of true and false belief ascription. *Cognition, 115*, 54-70. http://doi.org/10.1016/j.cognition.2009.11.008

Bakker, M., Hartgerink, C. H. J., Wicherts, J. M., & van der Maas, H. L. J. (2016).

Researchers' intuitions about power in psychological research. *Psychological Science,*

27, 1069-1077. http://doi.org/10.1177/0956797616647519

Bal, P. M. & Veltkamp, M. (2013). How does fiction reading influence empathy? An

experimental investigation of the role of emotional transportation. *PLOS ONE, 8*, 1-12.

http://doi.org/10.1371/journal.pone.0055341

Bal, M. (2009). *Narratology: Introduction to the theory of narrative* (3rd ed.). Toronto,

Canada: University of Toronto Press. (Original work published 1985)

Barnes, J. L. (2012). Fiction, imagination and social cognition: Insights from autism. *Poetics,*

40, 299-316. https://doi.org/10.1016/j.poetic.2012.05.001

Barnes, J. L. (2018). Imaginary engagement, real-world effects: Fiction, emotion, and social

cognition. *Review of General Psychology, 22,* 125-134.

https://doi.org/10.1037/gpr0000124

Barnes, J. L., Lombardo, M. V., Wheelwright, S., & Baron-Cohen, S. (2009). Moral dilemmas

film task: A study of spontaneous narratives by individuals with autism spectrum

conditions. *Autism Research, 2*, 148-156. http://doi.org/10.1002/aur.79

Baron-Cohen, S., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another advanced test

of theory of mind: Evidence from very high-functioning adults with autism or Asperger

Syndrome. *Journal of Child Psychology and Psychiatry, 38,* 813-822.

http://doi.org/10.1111/j.1469-7610.1997.tb01599.x

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of

mind"? *Cognition, 21*, 37-46. http://doi.org/10.1016/0010-0277(85)90022-8

Baron-Cohen, S., O'Riordan, M., Stone, V., Jones, R., & Plaisted, K. (1999). Recognition of

faux pas by normally developing children and children with Asperger syndrome or

high-functioning autism. *Journal of Autism and Developmental Disorders, 29,* 407-418.

http://doi.org/10.1023/a:1023035012436

Baron-Cohen, S., Richler, J., Bisarya, D., Gurunathan, N., & Wheelwright, S. (2003). The systemizing quotient: An investigation of adults with Asperger syndrome or high-functioning autism, and normal sex differences. *Philosophical Transactions of the Royal Society, B (Biological Sciences), 358,* 361–374. https://doi.org/10.1098/rstb.2002.1206

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

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

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin J., & Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): Evidence from Asperger Syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31* 5-17. http://doi.org/10.1023/A:1005653411471

Barnes, J. L., & Bloom, P. (2014). Children's preference for social stories. *Developmental Psychology, 50,* 498-503. https://doi.org/10.1037/a0033613

Barras, C. (2013, October 7). *Reading literary fiction makes you a nicer person.* New Scientist. Retrieved from https://www.newscientist.com/article/dn24356-reading-literary-fiction-makes-you-a-nicer-person/

Barrett, L. F., Mesquita, B., & Gendron, M. (2011). Context in emotion perception. *Current Directions in Psychological Science, 20,* 286-290. http://doi.org/10.1177/0963721411422522

Barrett, L. F., & Niedenthal, P. M. (2004). Valence focus and the perception of facial affect. *Emotion, 4,* 266-274. http://doi.org/10.1037/1528-3542.4.3.266

Barsade, S. G. (2002). The ripple effect: Emotional contagion and its influence on group behavior. *Administrative Science Quarterly, 47.* 644-675. https://doi.org/10.2307/3094912

Barthes, R. (1974). *S/Z: An essay.* New York, NY: Hill and Wang.

Batson, C. D. (1987). Prosocial motivation: Is it ever truly altruistic? In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 20, pp. 65–122). New York, NY: Academic Press.

Batson, C. D. (2011). *Altruism in humans*. New York, NY: Oxford University Press.

Batson, C. D., Duncan, B. D., Ackerman, P., Buckley, T. & Birch, K. (1981). Is empathic emotion a source of altruistic motivation? *Journal of Personality and Social Psychology, 40,* 290-302. http://doi.org/10.1037//0022-3514.40.2.290

Batson, C. D., Fultz, J., & Schoenrade, P. A. (1987). Distress and empathy: Two qualitatively distinct vicarious emotions with different motivational consequences. *Journal of Personality, 55*, 19–39. https://doi.org/10.1111/j.1467-6494.1987.tb00426.x

Batson, C. D., Lishner, D., & Stocks, E. (2015). The empathy-altruism hypothesis. In D. Schroeder, & W. Graziano (Eds.), *The Oxford handbook of prosocial behavior.* New York, NY: Oxford University Press.

Batson, C. D., Polycarpou, M. P., Harmon-Jones, E., Imhoff, H. J., Mitchener, E. C., Bednar, L. L., Klein, T. R., & Highberger, L. (1997). Empathy and attitudes: Can feeling for a member of a stigmatized group improve feelings toward the group? *Journal of*

Personality and Social Psychology, 7, 105-118. https://doi.org/10.1037//0022-3514.72.1.105

Batson, C. D., & Shaw, L. (1991). Evidence for altruism: Toward a pluralism of prosocial motives. *Psychological Inquiry, 2,* 107-122. https://doi.org/10.1207/s15327965pli0202_1

Baumeister, R. F., Bratlavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology, 5,* 323-370. https://doi.org/10.1037//1089-2680.5.4.323

Bazalgette, P. (2017). *The empathy instinct: How to create a more civil society*. London, England: John Murray.

Bazin, N., Brunet-Gouet, E., Bourdet, C., Kayser, N., Falissard, B., Hardy-Baylé, M., & Passerieux, C. (2009). Quantitative assessment of attribution of intentions to others in schizophrenia using an ecological video-based task: A comparison with manic and depressed patients. *Psychiatry Research, 167*, 28-35. http://doi.org/10.1016/j.psychres.2007.12.010

Beecher Stowe, H. (1995). *Uncle Tom's Cabin.* Hertfordshire, England: Wordsworth. (Original work published 1852)

Belluck, P. (2013, October 3). For better social skills, scientists recommend a little Chekhov. Retrieved from https://well.blogs.nytimes.com/2013/10/03/i-know-how-youre-feeling-i-read-chekhov

Biele, C., & Grabowska, A. (2006). Sex differences in perception of emotion intensity in dynamic and static facial expressions. *Experimental Brain Research 171*, 1-6. http://doi.org/10.1007/s00221-005-0254-0

Birch, S. A. J., & Bloom, P. (2007). The curse of knowledge in reasoning about false beliefs. *Psychological Science, 18*, 382-386. http://doi.org/10.1111/j.1467-9280.2007.01909.x

Black, J. E. (2018). An IRT analysis of the Reading the Mind in the Eyes Test. *Journal of Personality Assessment, 101,* 425-433. http://doi.org/10.1080/00223891.2018.1447946

Black, J. E., & Barnes, J. L. (2015a). Fiction and social cognition: The effect of viewing award-winning television dramas on theory of mind. *Psychology of Aesthetics, Creativity, and the Arts, 9*, 355-494. http://doi.org/10.1037/aca0000031

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

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

Bloom, P. (2016). *Against empathy: The case for rational compassion*. London, England: Penguin Random House.

Bloom, P. (2017). Empathy schmempathy: Response to Zaki. *Trends in Cognitive Sciences, 21,* 60-61. https://doi.org/10.1016/j.tics.2016.12.003

Bloom, P., & German, T. P. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition, 77*, 25-31. http://doi.org/10.1016/S0010-0277(00)00096-2

Bluck, S., Baron, J. M., Ainsworth, S. A., Gesselman, A. N., & Gold, K. L. (2013). Eliciting empathy for adults in chronic pain through autobiographical memory sharing. *Applied Cognitive Psychology*, *27,* 81–91. http://doi.org/10.1002/acp.2875

Bormann, D., & Greitemeyer, T. (2015). Immersed in virtual worlds and minds: Effects of in-game storytelling on immersion, need satisfaction, and affective theory of mind. *Social Psychological and Personality Science, 6*, 646–652. http://doi.org/10.1177/1948550615578177

Botvinick, M., Jha, A. P., Fabian, S. A., Solomon, P. E., & Prkachin, K. M. (2005). Viewing

facial expressions of pain engages cortical areas involved in the direct experience of

pain. *Neuroimage, 25,* 312-319. https://doi.org/10.1016/j.neuroimage.2004.11.043

Box, G. E. P., & Tidwell, P. W. (1962). Transformation of the independent variables.

Technometrics, 4, 531-550. http://doi.org/10.2307/1266288

Brecht, B. (1993). *The Resistible Rise of Arturo Ui* (S. S. Brecht, Trans.). Reading, England:

Methuen. (Original work published 1941)

Breithaupt, F. (2019). *The dark sides of empathy* (A. B. B. Hamilton, Trans.). Ithaca, NY:

Cornell University Press.

Brodie, R. (2009). *Virus of the mind: The revolutionary new science of the meme and how it

can help you.* London, England: Hay House.

Brown, D. E. (1991). *Human universals*. New York, NY: McGraw-Hill.

Bruner, J. (1986). *Actual minds, possible worlds.* Cambridge, MA: Harvard University Press.

Bryant, B. K. (1982). An index of empathy for children and adolescents. *Child Development,

53*, 413-425. http://doi.org/10.2307/1128984

Bull, R., Phillips, L. A., & Conway, C. A. (2008). The role of control functions in

mentalizing: Dual-task studies of theory of mind and executive function. *Cognition,

107*, 663-672. http://doi.org/10.1016/j.cognition.2007.07.015

Bureau of Labor Statistics (2017). *American time use survey – 2016 results* [PDF file].

Retrieved from https://www.bls.gov/news.release/archives/atus_06272017.pdf

Busselle, R., & Bilandzic, H. (2008). Fictionality and perceived realism in experiencing

stories: A model of narrative comprehension and engagement. *Communication Theory,

18*, 255-280. http://doi.org/10.1111/j.1468-2885.2008.00322.x

Busselle, R., & Bilandzic, H. (2009). Measuring narrative engagement. *Media Psychology,

12*, 321–347. http://doi.org/10.1080/15213260903287259

Cadwell, O. G. (2015). Literary fiction's influence on social cognitive brain activity. *Proceedings of the National Conference on Undergraduate Research* (pp. 570-574). Cheney, WA: Eastern Washington University.

Calame, C. (1999). *Performative aspects of the choral voice in Greek tragedy: Civic identity in performance.* In S. Goldhill & R. Osborne (Eds.), *Performance, culture and Athenian democracy* (pp. 125-153). Cambridge, England: Cambridge University Press.

Calvino, I. (1998). *If on a winter's night a traveller.* London, England: Vintage. (Original work published 1980)

Camerer, C. F. (2003). *Behavioral game theory: Experiments in strategic interaction*. Princeton, NJ: Princeton University Press.

Camerer, C. F. Dreber, A., Holzmeister, F., Ho, T-H., Huber, J., Johannesson, M., Kirchler, M., Navé, M., Nosek, B., Pfeiffer, T., Altmejd, A., Buttrick, N., Chan, T., Chen, Y., Forsell, E., Gampa, A., Heikensten, E., Hummer, L., Imai, T., … Wu, H. (2018). Evaluating the replicability of social science experiments in *Nature* and *Science* between 2010 and 2015. *Nature Human Behaviour, 2,* 637-644. http://doi.org/10.1038/s41562-018-0399-z

Castano, E. (2012). Anti-social behavior in individuals and groups: An empathy-focused approach. In K. Deux & M. Snyder (Eds.), *The Oxford Handbook of Personality and Social Psychology* (pp. 419-445). New York, NY: Oxford University Press.

Castelli, F., Happé, F., Frith, U., & Frith, C. (2000). Movement and mind: A functional imaging study of perception and interpretation of complex intentional movement patterns. *NeuroImage, 12,* 314–325. http://doi.org/10.1006/nimg.2000.0612

Chaiken, S. (1987). The heuristic model of persuasion. In M. P. Zanna, J. M. Olson, & C. P. Herman (Eds.), *Social Influence: The Ontario Symposium*, *5.* (pp. 3–39). Hillsdale, NJ: Erlbaum.

Chapman, J. (2015). *Emotionally durable design: Objects, experiences and empathy* (2nd ed.)*. London, England: Routledge.

Cheung, P. C., Ma, H. K., & Shek, T. L. D. (1998). Conceptions of success: Their correlates with prosocial orientation and behaviour in Chinese adolescents. *Journal of Adolescence*, *21*, 31-42. https://doi.org/10.1006/jado.1997.0127

Christensen, W. & Michael, J. (2016). From two systems to a multi-systems architecture for mindreading. *New Ideas in Psychology, 40,* 48-64. https://doi.org/10.1016/j.newideapsych.2015.01.003

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, J. (2001). Defining identification: A theoretical look at the identification of audiences with media characters. *Mass Communication and Society, 4*, 245-264. https://doi.org/10.1207/S15327825MCS0403_01

Coll, M., Viding, E., Rütgen, M., Silani, G., Lamm, C., Catmur, C. & Bird, G. (2017). Are we really measuring empathy? Proposal for a new measurement framework. *Neuroscience and Biobehavioral Reviews, 83,* 132-139. http://doi.org/10.1016/j.neubiorev.2017.10.009

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

Cook, R. D. & Weisberg, S. (1982). Criticism and influence analysis in regression. *Sociological Methodology, 13,* 313-361. https://doi.org/10.2307/270724

Coplan, A. (2004). Empathic engagement with narrative fictions. *Journal of Aesthetics and Art Criticism, 62,* 141-152. https://doi.org/10.1111/j.1540-594X.2004.00147.x

Coplan, A. (2011). Understanding empathy: Its features and effects. In A. Coplan & P. Goldie (Eds.), *Empathy: Philosophical and psychological perspectives* (pp. 3-18). Oxford, England: Oxford University Press.

Costa, P. T., & McCrae, R. R. (1992). *NEO PI-R professional manual*. Odessa, FL: Psychological Assessment Resources.

Cowley, M. (1958). *Writers at work.* New York, NY: The Viking Press.

Cox, C. L., Uddin, L. Q., Di Martino, A., Castellanos, F. X., Milham, M. P., & Kelly, C. (2012). The balance between feeling and knowing: Affective and cognitive empathy are reflected in the brain's intrinsic functional dynamics. *Social Cognitive and Affective Neuroscience, 7,* 727-737. https://doi.org/10.1093/scan/nsr051

Creese, B. (2015). *An assessment of the English and maths skills levels of prisoners in England.* London, England: UCL Institute of Education (Centre for Education in the Criminal Justice System).

Csibra, G. (2008). Goal attribution to inanimate agents by 6.5 month-old infants. *Cognition, 107,* 705-717. http://doi.org/10.1016/j.cognition.2007.08.001

Cuff, B. M. P., Brown, S. J., Taylor, L., & Howat, D. J. (2016). Empathy: A review of the concept. *Emotion Review, 7,* 144-153. https://doi.org/10.1177/1754073914558466

Currie, G. (2011). Empathy for objects. In A. Coplan & P. Goldie (Eds.), *Empathy: Philosophical and psychological perspectives* (pp. 82-98). Oxford, England: Oxford University Press.

Dalley, S. (Trans.). (2000). *Myths from Mesopotamia: Creation, the Flood, Gilgamesh, and Others.* New York, NY: Oxford University Press. (Original work published 1989)

Dalton, M. M., & Linder, L. R. (2005). *The sitcom reader: America viewed and skewed.* New York, NY: State University of New York Press.

Dan-Glauser, E. S., & Scherer, K. R. (2011). The Geneva affective picture database

(GAPED): A new 730-picture database focusing on valence and normative significance.

Behavior Research Methods, 43, 468-477. http://doi.org/10.3758/s13428-011-0064-1

David, S. (2016). *Eco-fiction: Bringing climate change into the imagination* (Doctoral thesis,

University of Exeter, Devon, England). Retrieved from

https://ore.exeter.ac.uk/repository/handle/10871/24331

Davies, M. (1993). The mental simulation debate. *Proceedings of the British Academy, 83,*

99-127. https://www.thebritishacademy.ac.uk/sites/default/files/83p099.pdf

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS*

Catalog of Selected Documents in Psychology, 10, 1-19.

http://www.uv.es/~friasnav/Davis_1980.pdf

Davis, M. (1983). Measuring individual differences in empathy: Evidence for a

multidimensional approach. *Journal of Personality and Social Psychology, 44,* 113-126.

http://doi.org/10.1037/0022-3514.44.1.113

Davis, M. H., Mitchell, K. V., Hall, J. A., Lothert, J., Snapp, T. & Meyer, M. (1999).

Empathy, expectations, and situational preferences: Personality influences on the

decision to participate in volunteer helping behaviors. *Personality, 67,* 469-503.

http://doi.org/10.1111/1467-6494.00062

De Graaf, A., Hoeken, H., Sanders, J., & Beentjes, W. J. (2012). Identification as a

mechanism of narrative persuasion. *Communication Research, 39,* 802-823.

https://doi.org/10.1177/0093650211408594

De Mulder, H. N. M., Hakemulder, F., van den Berghe, R., Klassen, F., & van Berkum, J. J.

A. (2017). Effects of exposure to literary narrative fiction: From book smart to street

smart? *Scientific Study of Literature, 7,* 129-169. https://doi.org/10.1075/ssol.7.1.06dem

De Vignemont, F. B. M., & Singer, T. (2006). The empathic brain: How, when and why? *Trends in Cognitive Sciences, 10,* 435-441. https://doi.org/10.1016/j.tics.2006.08.008

De Waal, F. (2011). A bottom-up view of empathy. In F. B. M. de Waal & Ferrari, P. F. (Eds.), *The primate mind* (pp. 121-138). Cambridge MA: Harvard University Press.

Dennett, D. (1978). Beliefs about beliefs. *Behavioral and Brain Sciences, 1*, 568-570. http://doi.org/10.1017/S0140525X00076664

Department for Business, Innovation and Skills. (2015). *2010 to 2015 government policy: public understanding of science and engineering* [Policy paper]. Retrieved from https://www.gov.uk/government/publications/2010-to-2015-government-policy-public-understanding-of-science-and-engineering/2010-to-2015-government-policy-public-understanding-of-science-and-engineering

Djikic, M. & Oatley, K. (2014). The art in fiction: From indirect communication to changes in the self. *Psychology of Aesthetics, Creativity and the Arts, 8*, 498-505. http://doi.org/10.1037/a0037999

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

Djikic, M., Oatley, K., Zoeterman, S., & Peterson, J. B. (2009). On being moved by art: How reading fiction transforms the self. *Creativity Research Journal, 21,* 24-29. https://doi.org/10.1080/10400410802633392

Dobs, K., Bülthoff, I., & Schulz, J. (2018). Use and usefulness of dynamic face stimuli for face perception studies–a review of behavioral findings and methodology. *Frontiers in Psychology,* 9, 1-7. https://doi.org/10.3389/fpsyg.2018.01355

Dodell-Feder, D., Lincoln, S. H., Coulson, J. P., & Hooker, C. I. (2013). Using fiction to

assess mental state understanding: A new task for assessing theory of mind in adults.

PLOS ONE, 8, 1-14. http://doi.org/10.1371/journal.pone.0081279

Dodell-Feder, D., Koster-Hale, J., Bedny, M., & Saxe, R. (2011). fMRI item analysis in a

theory of mind task. *Neuroimage, 55*, 705-712.

http://doi.org/10.1016/j.neuroimage.2010.12.040

Dodell-Feder, D., Ressler, K. J., & Germine, L. T. (2020). Social cognition or social class and

culture? On the interpretation of differences in social cognitive performance.

Psychological Medicine, 50, 133-145. https://doi.org/10.1017/S003329171800404X

Dodell-Feder, D. & Tamir, D. I. (2018). Fiction reading has a small positive impact on social

cognition: A meta-analysis. *Journal of Experimental Psychology: General*, *147,* 1713-

1727. http://doi.org/10.1037/xge0000395

Does, S., Ellemers, N., Dovido, J. F., Norman, J. B., Mentovich, A., van der Lee, R., & Goff,

P. A. (2018). Implications of research staff demographics for psychological science.

American Psychologist, 73, 639-650. https://doi.org/10.1037/amp0000199

Doherty, R. W. (1997). The Emotional contagion scale: A measure of individual differences.

Journal of Nonverbal Behavior, 21, 131-154. https://doi.org/10.1023/A:1024956003661

Dunn, R. (2012, March 2). *The story of the most common bird in the world.* Smithsonian

Magazine. Retrieved from www.smithsonianmag.com/science-nature/The-Story-of-the-

Most-Common-Bird-in-the-World.html

Duval, C., Piolino, P., Bejanin, A., Eustache, F., & Desgranges, B. (2010). Age effects on

different components of theory of mind. *Consciousness and Cognition 20,* 627-642.

http://doi.org/10.1016/j.concog.2010.10.025

Dziobek, I., Fleck, S., Kalbe, E., Rogers, K., Hassenstab, J., Brand, Kessler, J., Woike, J. K.,

Wolf, O. T., & Convit, A. (2006). Introducing MASC: A movie for the assessment of

social cognition. *Journal of Autism Developmental Disorders*, *36*, 623-636. http://doi.org/10.1007/s10803-006-0107-0

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

Eisenberg, N., & Strayer, J. (Eds.). (1987). *Cambridge studies in social and emotional development. Empathy and its development.* Cambridge, England: Cambridge University Press.

Ekman, P., & Friesen, W. V. (1971). Constants across culture in the face and emotion. *Journal of Personality and Social Psychology, 17,* 124-29. http://doi.org/10.1037/h0030377

Ekman, P., & Friesen, W. V. (1974). Nonverbal behavior and psychopathology. In R. J. Friedman & M. M. Katz (Eds.), *The Psychology of Depression* (pp. 203-224). Washington, DC: Winston & Sons. https://www.paulekman.com/wp-content/uploads/2013/07/Nonverbal-Behavior-And-Psychopathology.pdf

Ekman, P. & Friesen, W. V. (1976). *Pictures of facial affect.* Consulting Psychologists Press, Palo Alto, CA.

EmpathyLab (2019). *Pioneer schools: Evaluation report 2016-2018 executive summary* [PDF file]. Retrieved from https://irp-cdn.multiscreensite.com/b2f3fbc2/files/uploaded/Executive%20Summary%20of%20EmpathyLab%20Report%2C%20schools%20trial%2016-18.pdf

EmpathyLab (2017, January 31). *EmpathyLab at Moorlands Primary Academy*. Retrieved from http://www.empathylab.uk/empathylab-at-moorlands-primary-academy

Epley, N., Morowedge, C. K., & Keysar, B. (2004). Perspective taking in children and adults: Equivalent egocentricism but differential correction. *Journal of Experimental Social Psychology, 40*, 760-768. http://doi.org/10.1016/j.jesp.2004.02.002

Escalas, J. E. (2007). Self-referencing and persuasion: Narrative transportation versus analytical elaboration. *Journal of Consumer Research, 33,* 421-429. https://doi.org/10.1086/510216

Esterhuizen, T. M., & Thabane, L. (2016). Con: Meta-analysis: some key limitations and potential solutions. *Nephrology Dialysis Transplantation*, 31, 882–885. https://doi.org/10.1093/ndt/gfw092

Faber, M. (2000). *Under the skin.* Edinburgh, Scotland: Canongate.

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,* 175-191. https://doi.org/10.3758/bf03193146

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

Field, S., Meyer, A., & Witte, G. (2001). *Drehbuchschreiben fur fersehen und film*. Munchen: Ullstein Verlag.

Fiske, S. T., & Taylor, S. E. (2013). Social cognition: From brains to culture. Thousand Oaks, CA: Sage publications.

Flesch, R. (1948). A new readability yardstick. *Applied Psychology, 32*, 221-233. http://doi.org/10.1037/h0057532

Flynn, G. (2012). *Gone girl*. New York, NY: Crown.

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*, 370-376. http://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,* 274-285. https://doi.org/10.1037/a0038864

Forster, E. M. (1974) *Aspects of the novel.* London, England: Penguin. (Original work published 1927)

Frank, M. G., & Svetieva, E. (2015). Microexpressions and deception. In M. K. Mandal, & A. Awasthi (Eds.), *Understanding facial expressions in communication*, [Springer: EBL version], pp. 227-242. http://doi.org/10.1007/978-81-322-1934-7_11

Fraser, A., Padilla-Walker, L. M., Coyne, S., Nelson, L. J., & Stockdale, L. (2012). Associations between violent video gaming, empathic concern, and prosocial behavior toward strangers, friends, and family members. *Journal of Youth and Adolescence, 41,* 636-649. http://doi.org/10.1007/s10964-012-9742-2

Freedberg, D., & Gallese, V. (2007). Motion, emotion and empathy in esthetic experience. *Trends in Cognitive Science, 11,* 197-203. http://doi.org/10.1016/j.tics.2007.02.003

Freedman, M., Binns, M. A., Black, S. E., Murphy, C., & Stuss, D. T. (2013). Theory of mind and recognition of facial emotion in dementia: Challenge to current concepts. *Alzheimer Disease and Associated Disorders, 27,* 56-61. http://doi.org/10.1097/WAD.0b013e31824ea5db

Frith, C.D., & Corcoran, R. (1996). Exploring "theory of mind" in people with schizophrenia. *Psychological Medicine, 26*, 521-530. http://doi.org/10.1017/S0033291700035601

Frost, R. (2013). *The collected poems of Robert Frost.* London, England: Vintage. (Original work published 1969)

Frischmann, T. B., Mauloua, M., & Procci, K. (2015). 3-D gaming environment preferences: Inversion of the *Y*-axis. *Ergonomics, 58,* 1792-1799. http://doi.org/10.1080/00140139.2015.1044573

Funk, J. B., Fox, C. M., Chan, M., & Curtiss, K. (2008). The development of the Children's Empathic Attitudes Questionnaire using classical and Rasch analyses. *Journal of Applied Developmental Psychology*, *29*, 187-196. http://doi.org/10.1016/j.appdev.2008.02.005

Gabriel, S., & Young, A. F. (2011). Becoming a vampire without being bitten: The narrative collective-assimilation hypothesis. *Psychological Science, 22*, 990-994. http://doi.org/10.1177/0956797611415541

Garrido, M. V., & Prada, M. (2017). KDEF-PT: Valence, emotional intensity, familiarity and attractiveness ratings of angry, neutral, and happy faces. *Frontiers in Psychology, 8,* 1-9. https://doi.org/10.3389/fpsyg.2017.02181

Gallagher, S., & Gallagher, J. (2019). Acting oneself as another: An actor's empathy for her character. *Topoi: An International Review of Philosophy, Online First,* 1-12. https://doi.org/10.1007/s11245-018-9624-7

Gavaler, C., & Johnson, D. R. (2017). The genre effect: A science fiction (vs. realism) manipulation decreases inference effort, reading comprehension, and perceptions of literary merit. *Scientific Study of Literature, 7*, 79-108. http://doi.org/10.1075/ssol.7.1.04gav

Gentile, D. A., Anderson, C. A., Yukawa, S., Ihori, N., Saleem, M., Ming, L. K., Shibuya, A., Liau, A. K., Khoo, A., Bushman, B. J., Huesmann, L. R., & Sakamoto, A. (2009). The effects of prosocial video games on prosocial behaviors: International evidence from correlational, longitudinal and experimental studies. *Journal of Personality and Social Psychology Bulletin, 35*, 752-763. http://doi.org/10.1177/0146167209333045

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition, 56,* 165–193. https://doi.org/10.1016/0010-0277(95)00661-H

Gerrig, R. (1993). *Experiencing narrative worlds. On the psychological activities of reading.* New Haven, CT: Yale University Press.

Gibbons, J. D., & Chakraborti, S. (2011). *Nonparametric statistical inference* (5th ed.). Boca Raton, FL: Chapman & Hall/CRC.

Gilbert, D. T. (1991). How mental systems believe. *American Psychologist, 46,* 107-119. https://doi.org/10.1037/0003-066X.46.2.107

Gilbert, D. T. & Hixon, J. G. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology, 60,* 509-517. http://doi.org/10.1037/0022-3514.60.4.509

Gilbert, D. T., Killingsworth, M. A., Eyre, R. N., & Wilson, T. D. (2009). The surprising power of neighborly advice. *Science, 323,* 1617–1619. http://doi.org/10.1126/science.1166632

Giuliana, I., & Carvalho, H. C. (2016). Emotional contagion and socialization: Reflection on virtual interaction. In S. Y. Tettegah & D. L. Espelage (Eds.). *Emotions, Technology and Behaviors* (pp 63-82). https://doi.org/10.1016/B978-0-12-801873-6.00004-2

Golan, O., Baron-Cohen, S., & Golan, Y. (2008). The "Reading the Mind in Films" task [Child Version]: Complex emotion and mental state recognition in children with and without autism spectrum conditions. *Journal of Autism and Developmental Disorders, 38*, 1534-1541. http://doi.org/10.1007/s10803-007-0533-7

Goldman, A. (1995). Empathy, mind and morals. In M. Davies & S. Stone (Eds.), *Mental simulation: Philosophical and psychological essays* (pp. 185-208) Oxford, England: Blackwell.

Goldman, A. I. (2006). *Simulating minds. The philosophy, psychology and neuroscience of mindreading.* Oxford, England: Oxford University Press. http://doi.org/10.1093/0195138929.001.0001

Goldman, A. I. (2011). Two routes to empathy: Insights from cognitive neuroscience. In A. Coplan & P. Goldie (Eds.), *Empathy: Philosophical and psychological perspectives* (pp. 31–44). Oxford, England: Oxford University Press. http://doi.org/10.1093/acprof:oso/9780199539956.001.0001

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

Goldstein, T., & Filipe, A. (2018). The interpreted mind: Understanding acting. *Review of General Psychology, 22,* 220-229. https://doi.org/10.1037/gpr0000116

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

Goldstein, T., Wu, K., & Winner, E. (2009). Actors are skilled in theory of mind but not empathy. *Imagination, Cognition and Personality, 29*, 115-133. http://doi.org/10.2190/IC.29.2.c

Gordon, R. M. (1996). *"Radical" simulation.* In P. Carruthers & P. K. Smith (Eds.), *Theories of theories of mind* (pp. 11-21). Cambridge, England: Cambridge University Press. https://doi.org/10.1017/CBO9780511597985.003

Grandin, T., & Johnson, C. (2009). *Animals make us human: Creating the best life for animals.* Boston, MA: Mariner Books.

Gray, K., Jenkins, A. C., Heberlein, A.S., & Wegner, D. M. (2010). Distortions of mind perception in psychopathology. *Proceedings of the National Academy of Sciences, 108*, 477–479. http://doi.org/10.1073/pnas.1015493108

Green, M. C. & Brock, T. C. (2000). The role of transportation in the persuasiveness of public

narratives. *Journal of Personality and Social Psychology, 79*, 701-721.

http://doi.org/10.1037//0022-3514.79.5.701

Green, M. C. & Brock, T. C. (2002). In the mind's eye: Transportation-imagery model of

narrative persuasion. In M.C. Green, J. J. Strange, & T. C. Brock (Eds.), *Narrative*

impact: Social and cognitive foundations (pp. 315-341). New York, NY: Psychology

Press.

Green, M.C., Garst, J. & Brock, T.C. (2004). The power of fiction: Determinants and

boundaries. In L. J. Shrum (Ed.) *Blurring the lines between entertainment and*

persuasion (pp. 161-176). Mahwah, NJ: Lawrence Earlbaum Associates.

Gregory, C., Lough, S., Stone, V., Erzinclioglu, S., Martin, L., Baron-Cohen, S. & Hodges. J.

R. (2002). Theory of mind in patients with frontal variant frontotemporal dementia and

Alzheimer's disease: Theoretical and practical implications. *Brain, 125*, 752-764.

http://doi.org/10.1093/brain/awf079

Griffiths, P. (1997). *What emotions really are: The problem of psychological categories.*

Chicago, London: University of Chicago Press.

Guastella, A. J., Hermens, D. F., Van Zweiten, A., Naismith, S. L., Lee, R. S. C., Cacciotti-

Saija, C., Scott, E. M., & Hickie, I. B. (2013). Social cognitive performance as a marker

of positive psychotic symptoms in young people seeking help for mental health

problems. *Schizophrenia Research, 149*, 77-82.

http://doi.org/10.1016/j.schres.2013.06.006

Guillory, J., Spiegel, J., Drislane, M., Weiss, B. P., Donner, W., & Hancock, J. T. (2011).

Upset now?: Emotion contagion in distributed groups. *Proceedings of the SIGCHI*

Conference on Human Factors in Computer Systems (pp. 745-748).

https://doi.org/10.1145/1978942.1979049

Hakemulder, J. (2000). *The Moral Laboratory: Experiments examining the effects of reading literature on social perception and moral self-concept*. Amsterdam, The Netherlands: John Benjamins.

Hakemulder, J. F. (2004). Foregrounding and its effect on readers' perception. *Discourse Processes, 38,* 193-218. https://doi.org/10.1207/s15326950dp3802_3

Halberstadt, A.G., Dennis, P.A., & Hess, U. (2011). The influence of family expressiveness, individuals' own emotionality, and self-expressiveness on perceptions of others' facial expressions. *Journal of Nonverbal Behavior, 35*, 35-50. http://doi.org/10.1007/s10919-010-0099-5

Hall, A. E., & Bracken, C. C. (2011). "I really liked that movie": Testing the relationship between trait empathy, transportation, perceived realism and movie enjoyment. *Media Psychology, 23,* 90-99. http://doi.org/10.1027/1864-1105/a000036

Hall, J. A. (1990). *Nonverbal sex differences: Accuracy of communication and expressive style.* Baltimore, MD: Johns Hopkins University Press.

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature, 450,* 557-559. http://doi.org/10.1038/nature06288

Hansen, C. H., & Hansen, R. D. (1988). Finding the face in the crowd: An anger superiority effect. *Journal of Personality and Social Psychology, 54,* 917-924. http://doi.org/10.1037/0022-3514.54.6.917

Happé, F., Brownell, H., & Winner, E. (1999). Acquired "theory of mind" impairments following stroke. *Cognition, 70*, 211-240. http://doi.org/10.1016/S0010-0277(99)00005-0

Happé, F. G. E. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Autism and Developmental Disorders, 24*, 129-154. http://doi.org/10.1007/BF02172093

Happé, F. G. E. (1995). The role of age and verbal ability in the theory of mind task

performance of subjects with autism. *Child Development, 66,* 843-855.

http://doi.org/10.1111/j.1467-8624.1995.tb00909.x

Happé, F. G. E., Winner, E., & Brownell, H. (1998). The getting of wisdom: Theory of mind

in old age. *Developmental Psychology, 34*, 358–362. http://doi.org/10.1037/0012-

1649.34.2.358

Harbaugh, W. T., Mayr, U., & Burghart, D. R. (2007). Neural responses to taxation and

voluntary giving reveal motives for charitable donations. *Science, 316,* 1622-1625.

https://doi.org/10.1126/science.1140738

Hartley, J., & Turvey, S. (2013). *What books can do behind bars. Report on the work of PRG

1999-2013* [PDF file]. Retrieved from

https://prisonreadinggroupscouk.files.wordpress.com/2016/03/what-books-can-do-

behind-bars.pdf

Harvey, A. (2016). *Funding arts and culture in a time of austerity* [PDF file]. Retrieved from

http://www.artscouncil.org.uk/sites/default/files/download-

file/Funding%20Arts%20and%20Culture%20in%20a%20time%20of%20Austerity%20

(Adrian%20Harvey).pdf

Hauser, D. J., & Schwarz, N. (2016). Attentive Turkers: MTurk participants perform better on

attention checks than do subject pool participants. *Behavior Research Methods, 48,* 400-

407. http://doi.org/10.3758/s13428-015-0578-z

Hsee, C. K., Hatfield, E., Carlson, J. G., & Chemtob, C. (1990). The effect of power on

susceptibility to emotion contagion. *Cognition and Emotion, 4,* 327-340.

https://doi.org/10.1080/02699939008408081

Heal, J. (1986) *Simulation, theory, and content.* In P. Carruthers & P. K. Smith (Eds.),

Theories of theories of mind (pp. 75-89). Cambridge, England: Cambridge University

Press.

Health Foundation (2016). *Person-centred care made simple* [PDF file]. Retrieved from

http://www.health.org.uk/sites/health/files/PersonCentredCareMadeSimple.pdf

Heavey, L., Phillips, W., Baron-Cohen, S., & Rutter, M. (2000). The Awkward Moments

Test: A naturalistic measure of social understanding in autism. *Journal of Autism

Developmental Disorders, 30*, 225-236. http://doi.org/10.1023/A:1005544518785

Heide, F. J., Porter, N., & Saito, P. K. (2012). Do you hear the people sing? Musical theatre

and attitude change. *Psychology of Aesthetics, Creativity and the Arts, 6*, 224-230.

http://doi.org/10.1037/a0027574

Heider, F. & Simmel, M. (1944). An experimental study of apparent behavior. *American

Journal of Psychology, 57,* 243-259. https://doi.org/10.2307/1416950

Hein, G., Silani, G., Preuschoff, K., Batson, D. C., & Singer, T. (2010). Neural responses to

ingroup and outgroup members' suffering predict individual differences in costly

helping. *Neuron, 68,* 149-160. http://doi.org/10.1016/j.neuron.2010.09.003

Hemingway, E. (1988). The end of something. In Hemingway, E, *In our time* (pp. 42-46).

London, England: Vintage.

Higham, C. (1982). *Bette: A biography of Bette Davis.* London, England: Hodder &

Stoughton.

Hollywood Foreign Press Association (n.d.). *Winners and Nominees Best Motion Picture.*

Golden Globe Awards. Retrieved from http://www.goldenglobes.com/winners-

nominees/best-motion-picture-drama

Holt-Lunstad, J., Smith, T. B., & Layton, J. B. (2010). Social relationships and mortality risk: A meta-analytic review. *PLOS Medicine, 7,* 1-20. http://doi.org/10.1371/journal.pmed.1000316

Hoppenstand, G. (Ed.). (1998). *Popular fiction: An anthology.* New York, NY: Longman.

Hsu, C., Conrad, M., & Jacobs, A. M. (2014). Fiction Feelings in Harry Potter: haemodynamic response in the mid-cingulate cortex correlates with immersive reading experience. *Neuroport, 25*, 1356-1361. http://doi.org/10.1097/WNR.0000000000000272

Hume, D. (1985). *A treatise of human nature.* Middlesex, England: Penguin. (Original work published 1739-40)

Humphrey, K., Underwood, G., & Lambert, T. (2012). Salience of the lambs: A test of the saliency map hypothesis with pictures of emotive objects. *Journal of Vision, 12,* 1-15. http://doi.org/10.1167/12.1.22

IBM Corp. (Released 2015). IBM SPSS Statistics for Macintosh, Version 23.0. Armonk, NY: IBM Corp.

IBM Corp. (Released 2016). IBM SPSS Statistics for Macintosh, Version 24.0. Armonk, NY: IBM Corp.

Ickes, W. (Ed.). (1997). *Empathic accuracy.* New York, NY: Guilford Press.

Ickes, W., Stinson, L., Bissonnette, V., & Garcia, S. (1990). Naturalistic social cognition: Empathic accuracy in mixed-sex dyads. *Journal of Personality and Social Psychology, 59*, 730-742. http://doi.org/10.1037/0022-3514.59.4.730

Iguarta, J.-J. (2010). Identification with characters and narrative persuasion through fictional feature films. *Communications, 35,* 347-373. http://doi.org/10.1515/comm.2010.019

Iguarta, J.-J., Barrios, I., & Ortega, F. (2012). Analysis of the image of immigration in prime television fiction. *Communication and Society, 2,* 5-28.

https://revistas.unav.edu/index.php/communication-and-

society/article/view/36162/30669

Inkei, P. (2019, March). *Public funding of culture in Europe* [PDF file]. Retrieved from

http://www.budobs.org/files/Public%20cultural%20funding%202004-2017.pdf

John, O. P., Donahue, E. M., & Kentle, R. L. (1991). The Big Five Inventory – Versions 4a

and 54. Berkley, CA: University of California.

https://www.ocf.berkeley.edu/~johnlab/bfi.htm

Johnson, D. R. (2012). Transportation into a story increases empathy, prosocial behavior, and

perceptual bias toward fearful expressions. *Personality and Individual Differences, 52*,

150-155. http://doi.org/10.1016/j.paid.2011.10.005

Johnson, D. R., Cushman, G. K., Borden, L. A., & McCune, M. S. (2013). Potentiating

empathic growth: Generating imagery while reading fiction increases empathy and

prosocial behavior. *Psychology of Aesthetics, Creativity and the Arts, 7,* 306-312.

https://doi.org/10.1037/a0033261

Johnson, H., Keen, S., & Pritchard, D. (2011). *Unlocking value: The economic benefit of the

arts in criminal justice.* NPC. Retrieved from

http://www.thinknpc.org/publications/unlocking-value/

Jolliffe, D., & Farrington, D. P. (2006). Development and validation of the Basic Empathy

Scale. *Journal of Adolescence, 29,* 589-611.

https://doi.org/10.1016/j.adolescence.2005.08.010

Jordan, M. R., Amir, D., & Bloom, P. (2016). Are empathy and concern psychologically

distinct? *Emotion, 16,* 1107–1116. http://doi.org/10.1037/emo0000228

Joyce, R. (2012). *The unlikely pilgrimage of Harold Fry.* London, England: Random House.

Kaiser, H. (1974). An index of factorial simplicity. *Psychometrika 39,* 31–36.

https://doi.org/10.1007/BF02291575

Kalbe, E., Schlegel, M., Sack, A. T., Nowak, D. A., Dafotakis, M., Bangard, C., Brand, M., Shamay-Tsoory, S., Onur, O. A., & Kessler, J. (2010). Dissociating cognitive from affective theory of mind: A TMS study. *Cortex, 46*, 769-780. http://doi.org/10.1016/j.cortex.2009.07.010

Kandalaft, M. R., Didehbani, N., Krawczyk, D. C., Allen, T. T., & Chapman, S. B. (2013). Virtual reality social cognition training for young adults with high-functioning autism. *Journal of Autism and Developmental Disorders, 43*, 33-44. http://doi.org/10.1007/s10803-012-1544-6

Kanske, P., Böckler, A., Trautwein, F., & Singer, T. (2015). Dissecting the social brain: Introducing the EmpaToM to reveal distinct neural networks and brain-behavior relations for empathy and theory of mind. *NeuroImage 122,* 6-19. http://doi.org/10.1016/j.neuroimage.2015.07.082

Kaufman, W. (2006) *The civil war in American culture.* Wiltshire, England: Edinburgh University Press.

Kauschke, C., Bahn, D., Vesker, M., & Schwarzer, G. (2019). The role of emotional valence for the processing of facial and verbal stimuli–positivity or negativity bias? *Frontiers in Psychology, 10,* 1-15. https://doi.org/10.3389/fpsyg.2019.01654

Keen, S. (2007). *Empathy and the novel.* Oxford, England: Oxford University Press.

Keen, S. (2011). Reader's temperaments and fictional character. *New Literary History, 42,* 295-314. https://doi.org/10.1353/nlh.2011.0013

Keysers, C., Wicker, B., Gazzola, V., Anton, J.-L., Fogassi, L., & Gallese, V. (2004). A touching sight: SII/PV activation during the observation and experience of touch. *Neuron, 42,* 335-346. https://doi.org/10.1016/S0896-6273(04)00156-4

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

Kidd, D. C., & Castano, E. (2017a). Different stories: How levels of familiarity with literary

and genre-fiction relate to mentalizing. *Psychology of Aesthetics, Creativity and the*

Arts, 11, 474-486. http://doi.org/10.1037/aca0000069

Kidd, D. C., & Castano, E. (2017b). Panero et al. (2016): Failure to replicate methods caused

failure to replicate results. *Journal of Personality and Social Psychology, 112*, 1-4

http://doi.org/10.1037/pspa0000072

Kidd, D. C., & Castano, E. (2018a). Reading literary fiction can improve theory of mind.

Nature Human Behavior, 2, 604, https://doi.org/10.1038/s41562-018-0408-2

Kidd, D. C., & Castano, E. (2018b). Reading literary fiction and theory of mind: Three

preregistered replications and extensions of Kidd and Castano (2013). *Social*

Psychological and Personality Science, 20, 1-10.

http://doi.org/10.1177/1948550618775410

Kidd, D. C., Ongis, M., & Castano, E. (2016). On literary fiction and its effects on theory of

mind. *Scientific Study of Literature, 6*, 42-58. https://doi.org/10.1075/ssol.6.1.04kid

Kilts, C. D., Egan, G., Gideon, D. A., Ely, T. D., & Hoffman, M. (2003). Dissociable neural

pathways are involved in the recognition of emotion in static and dynamic facial

expressions. *NeuroImage, 18,* 156-168. http://doi.org/10.1006/nimg.2002.1323

Klimecki, O. M., Mayer, S. V., Jusyte, A., Scheeff, J., & Schönenberg, M. (2016). Empathy

promotes altruistic behavior in economic interactions. *Scientific Reports*, *6*, 1-5.

http://doi.org/10.1038/srep31961

Konrath, S. (2013). *Critical synthesis package: Interpersonal Reactivity Index (IRI)* [Critical

analysis, instrument and scoring guide]. MedEdPORTAL, 9(9596).

https://doi.org/10.15766/mep_2374-8265.9596

Konrath, S. H., O'Brien, E. H., & Hsing, C. (2011). Changes in dispositional empathy in

American college students over time: A meta-analysis. *Journal of Personality and

Social Psychology Review, 15,* 180-198. http://doi.org/10.1177/1088868310377395

Koolstra, C., Peeters, A., & Spinhof, H. (2002). The pros and cons of dubbing and subtitling.

European Journal of Communication, 17, 325-354.

https://doi.org/10.1177/0267323102017003694

Koopman, E. M. (2015). Empathic reactions after reading: The role of genre, personal factors

and affective responses. *Poetics, 50,* 62-79. http://doi.org/10.1016/j.poetic.2015.02.008

Koopman, E. M. (2016). Effects of "literariness" on emotions and on empathy and reflection

after reading. *Psychology of Aesthetics, Creativity and the Arts, 10*, 82-98.

http://doi.org/10.1037/aca0000041

Koopman, E. M. (2018). Does originality evoke understanding? The relation between literary

reading and empathy. *Review of General Psychology, 22,* 169-177.

http://dx.doi.org/10.1037/gpr0000107

Koopman, E. M. & Hakemulder, F. (2015). Effects of literature on empathy and self-

reflection: A theoretical-empirical framework. *Journal of Literary Theory, 9,* 79-111.

http://doi.org/10.1515/jlt-2015-0005

Kumkale, G. T., & Albarracín, D. (2004). The sleeper effect in persuasion: A meta-analytic

review. *Psychological Bulletin, 130,* 143-172. http://doi.org/10.1037/0033-

2909.130.1.143

Kurby, C. A., & Zacks, J. M. (2013). The activation of modality-specific representations

during discourse processing. *Brain & Language, 126,* 338-349.

http://doi.org/10.1016/j.bandl.2013.07.003

Kuzmičová, A., Mangen, A., Støle, H., & Begnum, A. C. (2017). Literature and readers' empathy: A qualitative text manipulation study. *Language and Literature: International Journal of Stylistics, 2,* 137-152. https://doi.org/10.1177/0963947017704729

Ladder to the Moon. (2015). *Personalised wellbeing.* Retrieved from http://www.laddertothemoon.co.uk/whyladder/

Lagerfeld, S. (1986). The reading revolution. *The Wilson Quarterly, 10*, 104-115. https://www.jstor.org/stable/40256998?seq=1

Lamott, A. (1995). *Bird by Bird: Some Instructions on Writing and Life*. New York: Anchor. (Original work published 1994).

Lanzoni, S. (2018). *Empathy: A history.* London, England: Yale University Press.

Larsen, N., Lee, K., & Ganea, P. A. (2017). Do storybooks with anthropomorphized animal characters promote prosocial behaviors in young children? *Developmental Science, 21*, 1-9. https://doi.org/10.1111/desc.12590

Le Bouc, R., Lenfant, P., Delbeuck, X., Ravasi, L., Lebert, F., Semah, F., & Pasquier, F. (2012). My belief or yours? Differential theory of mind deficits in frontotemporal dementia and Alzheimer's disease. *Brain, 135*, 3026-3038. http://doi.org/10.1093/brain/aws237

Lee, H. (1960). *To kill a mockingbird.* St. Ives, England: Heinemann New Windmills.

Lehne, M., Engel, P., Rohrmeier, M., Menninghaus, M., Jacobs, A. M., & Koelsch, S. (2015). Reading a suspenseful literary text activates brain areas related to social cognition and predictive inference. *PLOS ONE, 10*, 1-18. http://doi.org/10.1371/journal.pone.0124550

Leppänen, J. M., & Hietanen, J. K. (2003). Positive facial expressions are recognized faster than negative facial expressions, but why? *Psychological Research, 69,* 22-29. https://doi.org/10.1007/s00426-003-0157-2

Lindvall, A. M. C. (2008). Exploring the validity of the IPT-15: Using a multiple method approach to replicate and expand a peer-rating study of nonverbal communication. Retrieved from http://urn.nb.no/URN:NBN:no-19915

Lind, S.E., & Bowler, D.M. (2009). Language and theory of mind in autism spectrum disorder: The relationship between complement syntax and false belief task performance. *Journal of Autism and Developmental Disorders, 39*, 929-937. https://doi.org/10.1007/s10803-009-0702-y

Loughborough University (2015, April 14). *LISU's publishing statistics* [PDF file]. Retrieved from http://www.lboro.ac.uk/microsites/infosci/lisu/lisu-statistics/lisu-uk-library-statistics.pdf

Long, R., & Bolton, P. (2017). *English Baccalaureate and broadening the curriculum.* [House of Commons Library Briefing Paper]. Retrieved from https://researchbriefings.parliament.uk/ResearchBriefing/Summary/LLN-2017-0060

Lord, K. R., Lee, M-S., & Sauer, P. L. (1995). The combined influence hypothesis: Central and peripheral antecedents of attitude toward the ad. *Journal of Advertising*, *24*. 73–85. http://doi.org/10.1080/00913367.1995.10673469

Loveless, T. (2015). How well are American students learning? With sections on the gender gap in reading, effects of the Common Core, and student engagement. *The 2015 Brown Center Report on American Education, 3*. 1-38. https://www.brookings.edu/wp-content/uploads/2016/06/2015-Brown-Center-Report_FINAL-3.pdf

Lundqvist, L.-O., Carlsson, F., Hilmersson, P., & Juslin, P. N. (2009). Emotional responses to music: Experience, expression and physiology. *Psychology of Music, 37,* 61-90. https://doi.org/10.1177/0305735607086048

Lyons, V. & Gage, C. (2014). *Vibrant communities report: A proof of concept at Olive House* [PDF file]. Retrieved from

https://static1.squarespace.com/static/561235bde4b0e1ea75349256/t/56310548e4b051c e11cb9af2/1446053192926/Olive+House+report+2014.pdf

MacDorman, K. F. (2019). In the uncanny valley, transportation predicts narrative enjoyment more than empathy, but only for the tragic hero. *Computers in Human Behavior, 94,* 140-153. https://doi.org/10.1016/j.chb.2019.01.011

MacKenzie, S. B., Lutz, R. J., & Belch, G. E. (1986). The role of attitude toward the ad as a mediator of advertising effectiveness: A test of competing explanations. *Journal of Marketing Research, 23,* 130-143. http://doi.org/10.2307/3151660

Macrae, C.N., Johnston, L. (1998). Help, I need somebody: Automatic action and inaction. *Social Cognition, 16,* 400–417. https://doi.org/10.1521/soco.1998.16.4.400

Mangen, A., & Kuiken, D. (2014). Lost in an iPad: Narrative engagement on paper and tablet. *Scientific Study of Literature, 4*, 150-177. http://doi.org/10.1075/ssol.4.2.02man

Mann, C. (2011, November). *How the potato changed the world.* Smithsonian Magazine. Retrieved from https://www.smithsonianmag.com/history/how-the-potato-changed-the-world-108470605/

Marks, M., Chandler, L., & Baldwin, C. (2016). Environmental art as an innovative medium for environmental education in Biosphere Reserves. *Environmental Education Research, 23,* 1307-1321. https://doi.org/10.1080/13504622.2016.1214864

Mar, R. A. (2009). Empirical research on reading and watching narrative fiction. In D. Schram (Ed.). Reading and watching: *What does the written word have that images don't?* (pp. 53-61). Delft, The Netherlands: Eburon.

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

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,* 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,* 257-262. http://doi.org/10.1177/0963721417749654

Mar, R. A., & Oatley, K. (2008). The function of fiction is the abstraction and simulation of social experience. *Perspectives on Psychological Science, 3*, 173-192. http://doi.org/10.1111/j.1745-6924.2008.00073.x

Mar, R. A., Oatley, K., Djikic, M., & Mullin, J. (2011). Emotion and narrative fiction: Interactive influences before, during and after reading. *Cognition and Emotion, 25,* 818-833. http://doi.org/10.1080/02699931.2010.515151

Mar, R. A., Oatley, K., Hirsh, J., dela Paz, J. & Peterson, J.B. (2006). Bookworms versus nerds: Exposure to fiction versus non-fiction, divergent associations with social ability, and the simulation of fictional social worlds. *Research in Personality, 40,* 694-712. http://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,* 407-428. http://doi.org/10.1515/COMM.2009.025

Mar, R., Tackett, J. L., & Moore, C. (2010). Exposure to media and theory-of-mind development in preschoolers. *Cognitive Development, 25,* 69-78. http://doi.org/10.1016/j.cogdev.2009.11.002

Matsumoto, D., Olide, A., & Willingham, B. (2009). Is there an ingroup advantage in recognizing spontaneously expressed emotions? *Journal of Nonverbal Behavior, 33,* 181-191. http://doi.org/10.1007/s10919-009-0068-z

Matsumoto, D., LeRoux, J., Wilson-Cohn, C., Raroque, J., Kooken, K., Ekman, P., Yrizarry, N., Loewinger, S., Uchida, H., Yee, A., Amo, L., & Goh, A. (2000). A New test to measure emotion recognition ability: Matsumoto and Ekman's Japanese and Caucasian Brief Affect Recognition Test (JACBART). *Journal of Nonverbal Behavior, 24*, 179-209. https://doi.org/10.1023/A:1006668120583

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analysing data: A model comparison perspective* (2nd ed.). New York, NY: Psychology Press.

Maylor, E. A., Moulson, J. M., Muncer, A. M., & Taylor, L. A. (2002). Does performance on theory of mind tasks decline in old age? *British Journal of Psychology, 93,* 465-485. https://doi.org/10.1348/000712602761381358

Mazzocco, P. J., Green, M. C., Sasota, J. A., & Jones, N. W. (2010). This story is not for everyone: Transportability and narrative persuasion. *Social Psychological and Personality Science, 1*, 361-368. http://doi.org/10.1177/1948550610376600

McCall, C., & Singer, T. (2013). Empathy and the brain. In S. Baron-Cohen, H. Tager-Flusberg, & M. V. Lombardo (Eds.), *Understanding other minds: Perspectives from developmental social neuro-science* (pp. 195-213). Oxford, England: Oxford University Press. http://doi.org/10.1093/acprof:oso/9780199692972.003.0012

McDonald, D. G., Sarge, M. A., Lin, S.-F., Collier, J. G., & Potocki, B. (2015). A role for the self: Media content as triggers for involuntary autobiographical memories. *Communication Research, 42,* 3-29. http://doi.org/10.1177/0093650212464771

McDonald, S., Flanagan, S., Rollins, J., & Kinch, J. (2003). TASIT: A new clinical tool for assessing social perception after traumatic brain injury. *Journal of Head Trauma Rehabilitation, 18*, 219-238. http://doi.org/10.1097/00001199-200305000-00001

McGraw, K. O., & Wong, S. P. (1992). A common language effect size statistic. *Psychological Bulletin, 111,* 361-365. http://doi.org/10.1037/0033-2909.111.2.361

McLean, R. (Producer), & Slade, D. (Director). (2018). *Black mirror: Bandersnatch* [Motion picture]. United Kingdom: House of Tomorrow.

McLuhan, E. (n.d.). *Commonly asked questions (and answers).* The Estate of Corinne and Marshall McLuhan. Retrieved from https://marshallmcluhan.com/common-questions/

McLuhan, M. (1964/1994). *Understanding media: The extensions of man*. (Rev. ed.) MIT Press: Cambridge, MA.

McLuhan, M., Fiore, Q., & Agel, J. (2008). *The medium is the massage.* London, England: Penguin. (Original work published 1967)

McQuail, D. (2010). *Mass communication theory: An introduction.* London, England: Sage.

Miall, D. S., & Kuiken, D. (1994). Foregrounding, defamiliarization, and affect: Response to literary stories. *Poetics, 22,* 389–407. http://doi.org/10.1016/0304-422X(94)00011-5

Miall, D. S., & Kuiken, D. (1995). Aspects of literary response: A new questionnaire. *Research in the Teaching of English, 29,* 37-58. https://sites.ualberta.ca/~dmiall/MiallPub/Miall_Kuiken_LRQ_95.htm

Miesen, H. (2004). Fiction readers' appreciation of text attributes in literary and popular novels: Some empirical findings. *International Journal of Arts Management, 7,* 45-56. https://www.jstor.org/stable/41064830

Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and theory of mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child Development, 78,* 622-646. http://doi.org/10.1111/j.1467-8624.2007.01018.x

Mitchell, R. L. C., & Phillips, L. H. (2015). The overlapping relationship between emotion

perception and theory of mind. *Neuropsychologia, 70,* 1-10.

http://doi.org/10.1016/j.neuropsychologia.2015.02.018

Mitchell, P., Robinson, E. J., Isaacs, J. E., & Nye, R. M. (1996). Contamination in reasoning

about false belief: An instance of realist bias in adults but not children. *Cognition*, *59*, 1-

21. http://doi.org/10.1016/0010-0277(95)00683-4

Mol, S. E., & Bus, A. G. (2011). To read or not to read: A meta-analysis of print exposure

from infancy to early adulthood. *Psychological Bulletin, 137*, 267-296.

http://doi.org/10.1037/a0021890

Moore, M., & Gordon, P. C. (2015). Reading ability and print exposure: Item response theory

analysis of the author recognition test. *Behavior Research Methods, 47,* 1095–1109.

http://doi.org/10.3758/s13428-014-0534-3

Morgan, E. P. (1963). *Clearing the air.* New York, NY: Van Rees Press.

Morley, C. (2010). *Parnassus on wheels.* New York, NY: Melvin House. (Original work

published 1917)

Morreall, J. (2009). Enjoying incongruity. *Humor, 2,* 1-18.

https://doi.org/10.1515/humr.1989.2.1.1

MTM London. (2016). *Private investment in culture survey 2012/13, 2013/14 and 2014/15.*

[PDF file]. Retrieved from http://www.artscouncil.org.uk/sites/default/files/download-

file/Private_investment_culture_report_Nov_2016.pdf

Mumper, M. L., & Gerrig, R. J. (2017). Leisure reading and social cognition: A meta-

analysis. *Psychology of Aesthetics, Creativity and the Arts, 11*, 109-120.

http://doi.org/10.1037/aca0000089

Nabokov, V. (2000) *Lolita.* London, England: Penguin. (Original work published 1955)

National Theatre (n. d.). *The most significant plays of the 20th century*. Retrieved from http://www.nationaltheatre.org.uk/discover-more/platforms/nt2000-one-hundred-plays-of-the-century

Nell, V. (1988). The psychology of reading for pleasure: Needs and gratifications. *Reading Research Quarterly, 23*, 6-50. https://doi.org/10.2307/747903

Nettle, D. (2005). What Happens in Hamlet? Exploring the Psychological Foundations of Drama. In J. Gottschall & D. Sloan Wilson (Eds.), *The Literary Animal* (pp. 56-75). Evanston, Il: Northwestern University Press. http://doi.org/10.2307/j.ctvw1d5h1.10

Nettle. D. (2006). Psychological profiles of professional actors. *Personality and Individual Differences, 40,* 375-383. http://doi.org/10.1016/j.paid.2005.07.008

New York Theater (2013). *The 50 best plays of past 100 years.* Retrieved from http://newyorktheater.me/2013/06/27/the-50-best-plays-of-the-past-100-years-according-to-entertainment-weekly

Nichols, S., & Stich, S. P. (2003). *Mindreading: An integrated account of pretence, self-awareness, and understanding other minds.* Oxford, England: Clarendon Press.

Nomade, S., Genty, D., Sasco, R., Scao, V., Féruglio, V., Baffier, D., Guillou, H., Bourdier, C., Valladas, H., Reigner, E., Debard, E., Pastre, J.- F., & Geneste, J.-M. (2016). A 36,000-year-old volcanic eruption depicted in the Chauvet-Pont d'Arc cave (Ardèche, France)? *PLOS ONE, 11,* 1-10. http://doi.org/10.1371/journal.pone.0146621

Nomura, K., & Akai, S. (2012). Empathy with fictional stories: Reconsideration of the fantasy scale of the Interpersonal Reactivity Index. *Psychological Reports, 110,* 304-314. https://doi.org/10.2466/02.07.09.11.PR0.110.1.304-314

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Sciences Education, 15,* 625-632. http://doi.org/10.1007/s10459-010-9222-y

Nowicki, S. (2010). *A manual for diagnostic analysis of nonverbal accuracy 2*. Unpublished

manuscript. Department of Psychology, Emory University.

Nünning, V. (2014). *Reading fictions, changing minds: The cognitive value of fiction.*

Memmingen, Germany: Universitätsverlag Winter Heidelberg.

Nussbaum, M. (1990). *Love's knowledge: Essays on philosophy and literature.* New York,

NY: Oxford University Press.

Nussbaum, M. (1995). *Poetic justice: The literary imagination and public life*. Boston, MA:

Beacon.

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, 818-823. http://doi.org/10.1037/abn0000182

Oatley, K. (1994). A taxonomy of the emotions in literary response and a theory of

identification in fictional narrative. *Poetics, 23*, 53-74. https://doi.org/10.1016/0304-

422X(94)P4296-S

Oatley, K. (1999). Why fiction may be twice as true as fact: Fiction as cognitive and

emotional simulation. *Review of General Psychology, 3,* 101-117.

http://doi.org/10.1037/1089-2680.3.2.101

Oatley, K. (2011a, November). *Fiction hones social skills.* Scientific American. Retrieved

from https://www.scientificamerican.com/article/in-the-minds-of-others/

Oatley, K. (2011b). *Such stuff as dreams: The psychology of fiction.* Chichester, England:

Wiley-Blackwell.

Oatley, K. (2012). The cognitive science of fiction. *WIREs Cognitive Science, 3,* 425-430.

http://doi.org/10.1002/wcs.1185

Oatley, K. (2016). Fiction: Simulation of social worlds. *Trends in Cognitive Sciences, 20*,

618-628. http://doi.org/10.1016/j.tics.2016.06.002

Oatley, K. & Djikic, M. (2017). Psychology of narrative art. *Review of General Psychology, 22,* 161-168. https://doi.org/10.1037/gpr0000113

Office for National Statistics. (2012). *2011 Social Trends 41 - Lifestyles and social participation data* [Data set]. Retrieved from https://data.gov.uk/dataset/social_trends

Olejnik, S., & Algina, J. (2003). Generalized eta and omega squared statistics: Measures of effect size for some common research designs. *Psychological Methods, 8*, 434-447. https://doi.org/10.1037/1082-989X.8.4.434

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science, 308*, 255-258. http://doi.org/10.1126/science.1107621

Paal, T., & Bereczkei, T. (2007). Adult theory of mind, cooperation, Machiavellianism: The effect of mindreading on social relations. *Personality and Individual Differences, 43*, 541-551. http://doi.org/10.1016/j.paid.2006.12.021

Pallier, G., Wilkinson, R., Danthiir, V., Kleitman, S., Knezevic, G., Stankov, L., & Roberts, R. D. (2002). The role of individual differences in the accuracy of confidence judgments. *Journal of General Psychology 129,* 257-299. https://doi.org/10.1080/00221300209602099

Panero, M. E., Weisberg, D. S., Black, J., Goldstein, T. R., Barnes, J. L., Brownell, H., & Winner, E. (2016). Does reading a single passage of literary fiction really improve theory of mind? An attempt at replication. *Journal of Personality and Social Psychology, 111*, 46-54. http://doi.org/10.1037/pspa0000064

Panero, M. E., Weisberg, D. S., Black, J., Goldstein, T. R., Barnes, J. L., Brownell, H., & Winner, E. (2017). No support for the claim that literary fiction uniquely and immediately improves theory of mind: A reply to Kidd and Castano's commentary on Panero et al. (2016). *Journal of Personality and Social Psychology, 112*, 5-8. http://doi.org/10.1037/pspa0000079

Parsons, S., & Mitchell, P. (2002). The potential of virtual reality in social skills training for people with autistic spectrum disorders. *Journal of Intellectual Disability Research, 46*, 430-443. http://doi.org/10.1046/j.1365-2788.2002.00425.x

Pattengale, N., D., Alipour, M., Bininda-Emonds, O. R. P., Moret, B. M. E., & Stamatakis, A. (2009). How many bootstrap replicates are necessary? *Proceedings of the 13th Annual International Conference on Research in Computational Molecular Biology*, 184-200. https://doi.org/10.1007/978-3-642-02008-7_13

Penner, L. A., Fritzsche, B. A., Craiger, J. P., & Freifeld, T. R. (1995). Measuring the prosocial personality. In J. Butcher & C. D. Spielberger (Eds.), *Advances in personality assessment* (Vol 10) (pp. 147-163). Hillsdale, NJ: LEA.

Perner, J., & Ruffman, T. (2005). Infants' insight into the mind: How deep? *Science, 308*, 214-216. http://doi.org/10.1126/science.1111656

Perner, J., & Wimmer, H. (1985). "John thinks that Mary thinks that…" attribution of second-order beliefs by 5- to 10-year old children. *Experimental Child Psychology, 39,* 437-471. http://doi.org/10.1016/0022-0965(85)90051-7

Peterson, R. A., & Kim, Y. (2013). On the relationship between coefficient alpha and composite reliability. *Journal of Applied Psychology, 98,* 194-198. http://doi.org/10.1037/a0030767

Petro, N. M., Tong, T. T., Henley, D. J., & Neta, M. (2018). Individual differences in valence bias: fMRI evidence of the initial negativity hypothesis. *Social Cognitive and Affective Neuroscience, 13,* 687-698. https://doi.org/10.1093/scan/nsy049

Petty, R. E., & Cacioppo, J.T. (1981). *Attitudes and persuasion: Classic and contemporary approaches.* Dubuque, IA: Wm. C. Brown.

Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change.* New York, NY: Springer.

Pinker, S. (1997). *How the mind works.* London: Penguin.

Pinker, S. (2011). *The better angels of our nature: Why violence has declined.* New York, NY: Viking Adult.

Pino, M. C., & Mazza, M. (2016). The use of "literary fiction" to promote mentalizing ability. *PlOS ONE 11*, 1-14. http://doi.org/10.1371/journal.pone.0160254

Pinotti, A., & Salgaro, M. (2019). Empathy or empathies? Uncertainties in the interdisciplinary discussion. *Gestalt Theory, 41,* 141-158. http://doi.org/10.2478/gth-2019-0015

Poletti, M., Enrici, I., & Adenzato, M. (2012). Cognitive and affective theory of mind in neurodegenerative diseases: Neuropsychological, neuroanatomical and neurochemical levels. *Neuroscience and Biobehavioral Reviews, 36,* 2147-2164. http://doi.org/10.1016/j.neubiorev.2012.07.004

Pratchett, T. (1988). *Wyrd sisters.* London, England. Corgi Books.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40,* 879-891. http://doi.org/10.3758/BRM.40.3.879

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1*, 515-526. http://doi.org/10.1017/S0140525X00076512

Prescott, A. T., Sargent, J. D., & Hull, J. G. (2018). Metaanalysis of the relationship between violent video game play and physical aggression over time. *PNAS, 115,* 9882-9888. https://doi.org/10.1073/pnas.1611617114

Prison Reading Groups (2016). *Prison Reading Groups family days report 2016* [PDF file]. Retrieved from https://prisonreadinggroupscouk.files.wordpress.com/2016/03/prg-family-days-report-2016.pdf

Pritsch, C. Telkemeyer, S., Mühlenbeck, C., & Liebal, K. (2017). Perception of facial expressions reveals selective affect-based attention in humans and orangutans. *Scientific Reports, 7*, 1-12 http://doi.org/10.1038/s41598-017-07563-4

Radford, C. & Weston, M. (1975). How Can We Be Moved by the Fate of Anna Karenina? *Proceedings of the Aristotelian Society*, *49,* 67-93. http://www.jstor.org/stable/4106870

Rain, M., & Mar, R. (2014). Measuring reading behavior: Examining the predictive validity of print-exposure checklists. *Empirical Studies of the Arts, 32,* 93–108. http://doi.org/10.2190/EM.32.1f

Richter, D., Dietzel, C., & Kunzmann, U. (2010). Age differences in emotion recognition: The task matters. *The Journals of Gerontology, 66B,* 48-55. https://doi.org/10.1093/geronb/gbq068

Riedl, M. O., & Harrison, B. (2016). Using stories to teach human values to artificial agents. *Proceedings of the Workshops of the Thirtieth AAAI Conference on Artificial Intelligence AI, Ethics, and Society, 2016,* 105-112. http://www.aaai.org/Library/Workshops/ws16-02.php

Riggio, H. R., & Riggio, R. E. (2002). Emotional expressiveness, extraversion, and neuroticism: A meta-analysis. *Journal of Nonverbal Behavior, 26,* 195-218. http://doi.org/10.1023/A:1022117500440

Riva, P., & Andrighetto, L. (2012). "Everybody feels a broken bone but *only we* can feel a broken heart": Group membership influences the perception of targets' suffering. *European Journal of Social Psychology, 42,* 801-806. http://doi.org/10.1002/ejsp.1918

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research, 3,* 131-141. https://doi.org/10.1016/0926-6410(95)00038-0

Rocklage, M. D., Rucker, D. D., & Nordgren, L. F. (2018). Persuasion, emotion, and

language: The intent to persuade transforms language via emotionality. *Psychological

Science*, *29*, 749–760. https://doi.org/10.1177/0956797617744797

Rose, F. (2011). *The art of immersion: How the digital generation is remaking Hollywood.*

New York, NY: W. W. Norton & Company.

Rosenthal, R. (1976). *Experimenter effects in behavioral research.* New York, NY: John

Wiley.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological

Bulletin 86,* 638-641. http://doi.org/10.1037/0033-2909.86.3.638

Rosenthal, R. (2019). Parametric measures of effect size. In H. Cooper, & L. V. Hedges

(Eds.), *The handbook of research synthesis and meta-analysis* (pp. 231-244.). New

York, NY: Sage.

Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion.

Journal of Personality and Social Psychology Review, 5, 296-320.

https://doi.org/10.1207/S15327957PSPR0504_2

Rushdie, S. (1988) *The Satanic verses.* Reading, England: Vintage.

Rushton, J. P., Chrisjohn, R.D., & Fekken, G. C. (1981). The altruistic personality and the

self-report altruism scale. *Personality and Individual Differences, 1*, 292-302.

http://doi.org/10.1016/0191-8869(81)90084-2

Russell, J.A. (1980). A circumplex model of affect. *Journal of Personality and Social

Psychology, 39,* 1161-1178. https://doi.org/10.1037/h0077714

Rutherford, M.D., Baron-Cohen, S., & Wheelwright, S. (2002). Reading the mind in the

voice: A study with normal adults and adults with Asperger syndrome and high

functioning autism. *Journal of Autism and Developmental Disorders, 32*, 189-194.

http://doi.org/10.1023/A:1015497629971

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic

motivation, social development, and well-being. *American Psychologist, 55,* 68–78.

http://doi.org/10.1037/0003-066X.55.1.68

Samur, D., Tops, M., & Koole, S. L. (2018). Does a single session of reading literary fiction

prime enhanced mentalising performance? Four replication experiments of Kidd and

Castano (2013). *Cognition and Emotion, 32,* 130-144.

https://doi.org/10.1080/02699931.2017.1279591

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: The role of the

temporo-parietal junction in "theory of mind". *Neuroimage, 19*, 1835-1842.

https://doi.org/10.1016/S1053-8119(03)00230-1

Saxe, R., & Wexler, A. (2005). Making sense of another mind: The role of the right temporo-

parietal junction. *Neuropsychologia, 43*, 1391-1399.

https://doi.org/10.1016/j.neuropsychologia.2005.02.013

Schaafsma, S. M., Pfaff, D. W., Spunt, R. P., & Adolphs, R. (2015). Deconstructing and

reconstructing theory of mind. *Trends in Cognitive Sciences, 19*, 65-72.

https://doi.org/10.1016/j.tics.2014.11.007

Schellenberg, E. G. (2004). Music lessons enhance IQ. *Psychological Science, 15,* 511-514.

http://doi.org/10.1111/j.0956-7976.2004.00711.x

Schonert-Reichl, K. A., Smith, V., Zaidman-Zait, A., & Hertzman, C. (2012). Promoting

children's prosocial behaviors in school: Impact of the "Roots of Empathy" program on

the social and emotional competence of school-aged children. *School Mental Health, 4,*

1-21. http://doi.org/10.1007/s12310-011-9064-7

Schonfeld, Z. (2013, October 4). *Now We Have Proof Reading Literary Fiction Makes You A

Better Perso*n. The Atlantic. Retrieved from

https://www.theatlantic.com/culture/archive/2013/10/now-we-have-proof-reading-literary-fiction-makes-you-better-person/309996/

Schmid, P. C., & Schmid Mast, M. (2010). Mood effects on emotion recognition. *Motivation and Emotion, 34,* 288-292. http://doi.org/10.1007/s11031-010-9170-0

Seddon, C. (2011). *Lifestyles and social participation* [PDF file]. Retrieved from https://www.ons.gov.uk/ons/rel/social-trends-rd/social-trends/social-trends-41/lifestyles-chapter.pdf

Senju, A., Southgate, V., White, S., & Frith, U. (2009). Mindblind eyes: An absence of spontaneous theory of mind in Asperger syndrome. *Science, 325*(5942)*,* 883-885. http://doi.org/10.1126/science.1176170

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

Shim, H., Lim, S., Jung, E. E., & Shin, E. (2018). I hate binge-watching but I can't help doing it: The moderating effect of immediate gratification and need for cognition on binge-watching attitude-behavior relation. *Telematics and Informatics, 35,* 1971-1979. https://doi.org/10.1016/j.tele.2018.07.001

Siegal, M., & Beattie, K. (1991). Where to look first for children's knowledge of false beliefs. *Cognition, 38*(1), 1-12. https://doi.org/10.1016/0010-0277(91)90020-5

Simonsohn, U. (2005). Small telescopes: Detectability and the evaluation of replication results. *Psychological Science, 26,* 559-569. http://doi.org/10.1177/0956797614567341

Simonsohn, U., Nelson, L. D., & Simmons, J. P (2014). *P*-curve and effect size: Correcting for publication bias using only significant results. *Perspectives on Psychological Science, 9,* 666-681. https://doi.org/10.1177/1745691614553988

Singer, P. (2011). *The expanding circle: Ethics, evolution and moral progress.* Princeton, NJ:

Princeton University Press. (Original work published 1981)

Singer, T. (2006). The neuronal basis and ontogeny of empathy and mind reading: Review of

literature and implications for future research. *Neuroscience and Biobehavioral*

Reviews, 30, 855–863. http://doi.org/10.1016/j.neubiorev.2006.06.011

Singer, T., & Lamm, C. (2009). The social neuroscience of empathy. *New York Academy of*

Sciences, 1156, 81-96. http://doi.org/10.1111/j.1749-6632.2009.04418.x

Slater, M. D. (2002). Entertainment education and the persuasive impact of narratives. In

M.C. Green, J. J. Strange, & T. C. Brock (Eds.), *Narrative impact: Social and cognitive*

foundations (pp. 157-182). New York, NY: Psychology Press.

Slater, M. D., & Rouner, D. (2002). Entertainment-education and elaboration likelihood:

Understanding the processing of narrative persuasion. *Communication Theory, 12*, 173-

191. https://doi.org/10.1111/j.1468-2885.2002.tb00265.x

Speer, N. K., Reynolds, J. R., Swallow, K. M., & Zacks, J. M. (2009). Reading stories

activates neural representations of visual and motor experiences. *Psychological Science,*

20, 989-999. http://doi.org/10.1111/j.1467-9280.2009.02397.x

Spiers, H. J., & Maguire, E. A. (2006). Spontaneous mentalizing during an interactive real

world task: An fMRI study. *Neuropsychologia, 44,* 1674-1682.

http://doi.org/10.1016/j.neuropsychologia.2006.03.028

Spreng, R. N., McKinnon, M. C., Mar, R. A., & Levine, B. (2009). The Toronto Empathy

Questionnaire: Scale development and initial validation of a factor-analytic solution to

multiple empathy measures. *Journal of Personality Assessment, 91,* 62–71.

http://doi.org/10.1080/00223890802484381

Spunt, R. P., & Adolphs, R. (2014). Validating the why/how contrast for functional MRI

studies of theory of mind. *Neuroimage, 99,* 301-311.

http://doi.org/10.1016/j.neuroimage.2014.05.023

Smith, D., Schlaepfer, P., Major, K., Dyble, M., Page, A. E., Thompson, J., Chaudhary, N.,

Salali, G. D., Mace, R., Astete, L., Ngales, M., Vinicius, L., & Migliano, A. B. (2017).

Cooperation and the evolution of hunter-gatherer storytelling. *Nature Communications,*

8, 1-9. http://doi.org/10.1038/s41467-017-02036-8

Smith, J. (2017). What is empathy for? *Synthese, 194,* 709-722.

http://doi.org/10.1007/s11229-015-0771-8

Smith, Z. (2019, October 24). *Fascinated to presume: In defense of fiction.* The New York

Review of Books. Retrieved from https://www.nybooks.com/articles/2019/10/24/zadie-

smith-in-defense-of-fiction/

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,*

51–68. http://doi.org/10.3758/BF03208254

Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing.

Reading Research Quarterly, 24, 402-433. http://doi.org/10.2307/747605

Stanovich, K. E., West, R. F., & Harrison, M. R. (1995). Knowledge growth and maintenance

across the life span: The role of print exposure. *Developmental Psychology, 31,* 811–

826. http://doi.org/10.1037/0012-1649.31.5.811

Stansfield, J., & Bunce, L. (2014). The relationship between empathy and reading fiction:

Separate role for cognitive and affective components. *Journal of European Psychology*

Students, 5, 9-18. http://doi.org/10.5334/jeps.ca

Stephens-Hernandez, A. B., Livingston, J. N., Dacons-Brock, K., Craft, H. L., Cameron, A.,

Franklin, S. O., & Howlett, A. C. (2007). Drama-based education to motivate

participation in substance abuse prevention. *Substance Abuse Treatment Prevention and Policy, 2*, 1-11. http://doi.org/10.1186/1747-597X-2-11

Stoker, B. (1897/2019). *Dracula.* London, England: Vintage.

Stone, T., & Davies, M. (1996) *The mental simulation debate: A progress report* In P. Carruthers & P. K. Smith (Eds.), *Theories of theories of mind* (pp. 119-140). Cambridge, England: Cambridge University Press.

Storr, W. (2019). *The science of storytelling.* London, England: William Collins.

Sullivan, K., Zaitchik, D., & Tager-Flusberg, H. (1994). Preschoolers can attribute second-order beliefs. *Developmental Psychology, 30*, 395-402. http://doi.org/10.1037/0012-1649.30.3.395

Surguladze, S. A., Young, A. W., Senior, C., Brébion, G., Travis, M. J., & Phillips, M. L. (2004). Recognition accuracy and response bias to happy and sad facial expressions in patients with major depression. *Neuropsychology, 18,* 212-218. http://doi.org/10.1037/0894-4105.18.2.212

Sze, J. A., Goodkind, M.S., Gyurak, A., & Levenson, R.W. (2012). Aging and emotion recognition: Not just a losing matter. *Psychology and Ageing, 27*, 940-950. http://doi.org/10.1037/a0029367

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics.* (5th ed.). Boston, MA: Pearson Education.

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

Taylor, M., & Carlson, S. M. (1997). The relation between individual differences in fantasy and theory of mind. *Child Development, 68*, 436–455. https://www.jstor.org/stable/1131670?seq=1

Teding van Berkhout, E., & Malouff, J. M. (2016). The efficacy of empathy training: A meta-analysis of randomized controlled trials. *Counselling Psychology, 63,* 32-41. http://doi.org/10.1037/cou0000093

Titchener, E. B. (1909). *Experimental psychology of the thought-processes.* New York, NY: Macmillan.

The Reader's Digest Association (1950, August). [Freestanding quotation], Quote Page 80. *Reader's Digest* 57, 80.

Todd, R. M., Cunningham, W. A., Anderson, A. K., & Thompson, E. (2012). Affect-based attention as emotion regulation. *Trends in Cognitive Sciences, 16,* 365-372. https://doi.org/10.1016/j.tics.2012.06.003

Tormala, Z. L., & Petty, R. E. (2004). Source credibility and attitude certainty: A metacognitive analysis of resistance to persuasion. *Consumer Psychology*, *14*, 427–442. http://doi.org/10.1207/s15327663jcp1404_11

Unoka, Z., Fogd, D., Füzy, M., & Csukly, G. (2011). Misreading the facial signs: Specific impairments in error patterns in recognition of facial emotions with negative valence in borderline personality disorder. *Psychiatry Research, 189,* 419-425. http://doi.org/10.1016/j.psychres.2011.02.010

Valente, T. W., & Bharath, U. (1999). An evaluation of the use of drama to communicate HIV/AIDS information. *Education and Prevention, 11*, 203-211. https://www.ncbi.nlm.nih.gov/pubmed/10407454

Valkenburg, P. M., Peter, J., & Walther, J. B. (2016). Media effects: Theory and research. *Annual Review of Psychology, 67,* 315-338. http://doi.org/10.1146/annurev-psych-122414-033608

Van Baaren, R. B., Holland, R. W., Kawakami, K., & van Knippenberg, A. (2004). Mimicry and prosocial behavior. *Psychological Science, 15,* 71–74. http://doi.org/10.1111/j.0963-7214.2004.01501012.x

Van Kujik, I., Verkoeijen, P., 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: Psychology, 4*(1), 2-12. http://doi.org/10.1525/collabra.117

Van Laer, T., De Ruyter, K., Visconti, L. M. & Wetzels, M. (2014). The extended transportation-imagery model: A meta-analysis of the antecedents and consequences of consumers' narrative transportation. *Consumer Research, 40,* 797-817. http://doi.org/10.1086/673383

Van Peer, W., Hakemulder, J., & Zyngier, S. (2007). Lines on feeling: Foregrounding, aesthetics and meaning. *Language and Literature, 16,* 197–213. http://doi.org/10.1177/0963947007075985

Vischer, R. (1873). *Über das optische formgefühl: Ein beitrag zur ästhetik* [About the optical sense of form: A contribution to aesthetics]. Leipzig: Credner.

Von Trier, L. (Director), Jensen, P. A., & Windeløv, V. (Producers). (2000). *Dancer in the dark* [Motion picture]. Denmark: Zentropa.

Wai, M., & Tiliopoulos, N. (2012). The affective and cognitive empathic nature of the dark triad of personality. *Personality and Individual Differences, 52,* 794-799. http://doi.org/10.1016/j.paid.2012.01.008

Waite, M. (Ed.). (2012). Oxford English Dictionary (7th ed.). Oxford, England: Oxford University Press.

Wallentin, M., Nielsen, A. H., Vuust, P., Dohn, A., Roepstorff, A., & Lund, T. E. (2011). BOLD response to motion verbs in left posterior middle temporal gyrus during story

comprehension. *Brain & Language, 119,* 221-225.

http://doi.org/10.1016/j.bandl.2011.04.006

Walter, H. (2012). Social cognitive neuroscience of empathy: Concepts, circuits and genes.

Emotion Review, 4, 9-17. http://doi.org/10.1177/1754073911421379

Wang, J. J., Miletich, D. D., Ramsey, R., & Samson, D. (2014). Adults see vision to be more

informative than it is. *The Quarterly Journal of Experimental Psychology, 67*, 2279-

2292. http://doi.org/10.1080/17470218.2014.915331

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief

measures of positive and negative affect: The PANAS scales. *Journal of Personality

and Social Psychology, 54*, 1063-1070. http://doi.org/10.1037//0022-3514.54.6.1063

Weinfurt, K. P. (2000). Repeated measures analyses: ANOVA, MANOVA, and HLM. In L.

G. Grimm & P. R. Yarnold (Eds.), *Reading and Understanding more multivariate

statistics* (pp. 317-361). Washington, DC: American Psychological Association.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind

development: The truth about false belief. *Child Development, 72*, 655-684.

http://doi.org/10.1111/1467-8624.00304

West, R. F., Stanovich, K. E., & Mitchell, H. R. (1993). Reading in the real world and its

correlates. *Reading Research Quarterly, 28,* 34-50. https://doi.org/10.2307/747815

Whitman, W. (2017). *Leaves of grass.* Milton Keynes, England. (Original work published

1855)

Wicker, B., Keysers, C., Plailly, J., Royet, J. P., Gallese, V., & Rizzolatti, G. (2003). Both of

us disgusted in my insula: The common neural basis of seeing and feeling disgust.

Neuron, 40, 655-664. https://doi.org/10.1016/S0896-6273(03)00679-2

Willems, R. M., & Jacobs, A. M. (2016). Caring about Dostoyevsky: The untapped potential of studying literature. *Trends in Cognitive Sciences, 20,* 243-245. http://doi.org/10.1016/j.tics.2015.12.009

Winner, E., Brownell, H., Happé, F., Blum, A., & Pincus, D. (1998). Distinguishing lies from jokes: Theory of mind deficits and discourse interpretation in right hemisphere brain-damaged patients. *Brain and Language, 62,* 89–106. http://doi.org/10.1006/brln.1997.1889

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13,* 103-128. http://doi.org/10.1016/0010-0277(83)90004-5

Woronkowicz, J., Nichols, B., & Iyengar, S. (2012). *How the United States funds the arts.* [PDF file]. Retrieved from https://www.arts.gov/sites/default/files/how-the-us-funds-the-arts.pdf

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

Young, L., Bechara, A., Tranel, D., Damasio, H., Hauser, M., & Damasio, A. (2010). Damage to ventromedial prefrontal cortex impairs judgment of harmful intent. *Neuron, 65*, 845–51. https://doi.org/10.1016/j.neuron.2010.03.003

Zaki, J. (2017). Moving beyond stereotypes of empathy. *Trends in Cognitive Sciences, 21,* 59-60. https://doi.org/10.1016/j.tics.2016.12.004

Zaki, J. (2019). *The war for kindness: Building empathy in a fractured world.* London, England: Robinson.

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

Zemborain, M. R., & Johar, G. V. (2007). Attitudinal ambivalence and openness to

persuasion: A framework for interpersonal influence. *Journal of Consumer Research*,

33, 506–514. https://doi.org/10.1086/510224

Zhou, H., Majka, E. A., & Epley, N. (2017). Inferring perspective versus getting perspective:

Underestimating the value of being in another person's shoes. *Psychological Science,*

28, 482-493. http://doi.org/10.1177/0956797616687124

Zickfeld, J. H., Schubert, T. W., Siebt, B., & Fiske, A. P. (2017). Empathic concern is part of

a more general communal emotion. *Frontiers in Psychology, 8*, 1-16.

https://doi.org/10.3389/fpsyg.2017.00723

Zillman, D. (1988). Mood management through communication choices. *American*

Behavioral Scientist, 31, 327-340. Retrieved from

https://doi.org/10.1177%2F000276488031003005

Ziolkowski, T. (2011, December 8). *Gilgamesh among us: Modern encounters with the*

ancient epic. London, England: Cornell University Press.

Zipes, J. (2012). *The irresistible fairy tale: The cultural and social history of a genre.*

Princeton, NJ: Princeton University Press.

Zunshine, L. (2006). *Why we read fiction: Theory of mind and the novel*. Columbus, OH:

Ohio State University Press.

Zwaan, R. A. (2004). The immersed experiencer: Toward and embodied theory of language

comprehension. In B. H. Ross (Ed.), *The psychology of learning and motivation* (pp. 35-

62). New York, NY: Academic Press.

Appendix A

Fiction Media Exposure Test names and percentage of participants that selected each name. Names in the print-exposure scale were taken from *Behavior Research Methods, 40,* Acheson, D. J., Wells, J. B., & MacDonald, M. C., "New and updated tests of print-exposure and reading abilities in college students", 278-289. Copyright Psychonomic Society (2008), licensed by Creative Commons (CC BY-ND: https://creativecommons.org/licenses/by-nd/4.0/legalcode)

Scale	Name	% participants that identified name
Print-exposure	J. R. R. Tolkien	84
	George Orwell	82
	Ernest Hemingway	81
	T. S. Eliot	77
	Virginia Woolf	77
	F. Scott Fitzgerald	73
	Harper Lee	71
	Salman Rushdie	58
	James Joyce	57
	Tom Clancy	56
	J. D. Salinger	54
	James Patterson	53
	Margaret Atwood	53
	Maya Angelou	46
	John Irving	45
	Gabriel Garcia Marquez	33
	Kazuo Ishiguro	31
	Umberto Eco	30
	Kurt Vonnegut	27
	Alice Walker	27
	Thomas Wolfe	25
	Clive Cussler	24
	Anne McCaffrey	19
	Nora Ephron	17
	Saul Bellow	17
	Sue Grafton	14
	Jonathan Kellerman	12
	T. C. Boyle	9
	Ann Beattie	4
	Jane Smiley	3
Film-exposure	Avatar	93
	The Exorcist	92
	Gladiator	90
	Schindler's List	89
	Saving Private Ryan	88
	12 Years a Slave	88
	American Beauty	82
	Rain Man	80
	One Flew Over the Cuckoo's Nest	78
	A Beautiful Mind	77
	Spartacus	74
	Lawrence of Arabia	72

Appendix B

Supplemental analysis Study 1.

Partial correlations: Internal correlations remained significant among the IRI and FMET scales respectively, all ps < .001, and between altruism and all IRI scales, all ps < .01. Significant inter-correlations were as follows ($df = 393$): Fantasy remained associated with print-exposure, $r = .17$, $p = .001$, 95% CI [.07, .27], and correlations reached significance with play, $r = .15$, $p = .003$, 95% CI [.05, .24], and film-exposure, $r = .15$, $p = .002$, 95% CI [.05, .26]. Altruism remained associated with print-exposure, $r = .12$, $p = .016$, 95% CI [.01, .23], and play-exposure, $r = .15$, $p = .004$, 95% CI [.04, .25] but not film-exposure, $r = .07$, $p = .153$, 95% CI [-.03, .18].

Perspective-taking (PT) and empathic concern (EC) remained associated with ($df = 344$): preferences for romance, $r_{PT} = .12$, $p = .023$, 95% CI [.02, .23], $r_{EC} = .24$, $p < .001$, 95% CI [.13, .34], comedy, $r_{PT} = .14$, $p = .011$, 95% CI [.03, .24], $r_{EC} = .20$, $p < .001$, 95% CI [.09, .30], experimental/postmodern, $r_{PT} = .18$, $p = .001$, 95% CI [.07, .29], $r_{EC} = .17$, $p = .001$, 95% CI [.07, .27], and factual/documentary, $r_{PT} = .12$, $p = .025$, 95% CI [.02, .23] $r_{EC} = .12$, $p = .027$, 95% CI [.02, .23]; and empathic concern with drama, $r = .16$, $p = .003$ 95% CI [.05, .27]. Fantasy correlated with drama, $r = .28$, $p < .001$, 95% CI [.17, .38] romance, $r = .19$, $p < .001$, 95% CI [.09, .28], comedy, $r = .18$, $p = .001$, 95% CI [.08, .30] and experimental/postmodern, $r = .16$, $p = .002$, 95% CI [.06, .27]. Altruism correlated with comedy, $r = .18$, $p = .001$, 95% CI [.08, .28]. experimental/postmodern, $r = .16$, $p = .003$, 95% CI [.04, .27], and factual/documentary, $r = .12$, $p = .032$, 95% CI [.01, .22]. Fantasy remained associated with: ($df = 392$) reading novels, $r = .16$, $p = .001$, 95% CI [.06, .26], and the negative relationship with TV approached but did not reach significance, $r = -.10$, $p = .055$, 95% CI [-.19, -.001]. Altruism remained positively associated with plays, $r = .17$, $p = .001$, 95% CI [.07, .27] and negatively associated with TV, $r = -.17$, $p = .001$, 95% CI [-.27, -.08]

and film $r = -.11$, $p = .03$, 95% CI [-.21, -.01]. The correlation with reading novels did not reach significance, $r = .09$, $p = .093$, 95% CI [-.02, .20].

Media selection: Media selection is based on social, developmental and dispositional factors (for a review, see Valkenburg et al., 2016). Escapism was the primary motivation for engaging with fiction media (except with plays, which tended to be motivated by an interest in the creatives), followed by mood management (Zillman, 1988): people tended to select media to maintain or improve their mood-state (Figure 1).

Reading a novel because participants felt they "should" have read it negatively correlated with perspective-taking, $r_s (304) = -.18$, $p = .002$, 95% CI [-.28, -.08], and fantasy, $r_s (304) = -.19$, $p = .001$, 95% CI [-.30, .07], and watching a film because they "should" have seen it negatively correlated with empathic concern, $r_s (341) = -.11$, $p = .038$, 95% CI [-.22, .001], fantasy, $r_s (341) = -.12$, $p = .023$, 95% CI [-.23, -.02], and altruism, $r_s (341) = -.12$, $p = .029$, 95% CI [-.22, -.02]. In contrast, results showed positive correlations between watching a film to improve mood with empathic concern, $r_s (341) = .14$, $p = .012$, 95% CI [.04, .23], and using films for escapism with fantasy, $r_s (341) = .12$, $p = .029$, 95% CI [.01, .23]. Altruism was positively associated with wanting to learn about a topic through watching film, $r_s (341) = .14$, $p = .01$, 95% CI [.03, .25], TV, $r_s (295) = .18$, $p = .002$, 95% CI [.06, .29], and plays, $r_s (222) = .17$, $p = .009$, 95% CI [.04, .30], and reading a novel because of an interest in the writer, $r_s (304) = .12$, $p = .033$, 95% CI [.01, .23]. Correlations between reasons for media selection and empathic abilities suggest a mediating role for media presentation in relationships between the antecedents and consequences of fiction-engagement, which constitutes a central concept in several media effects theories (Valkenburg et al., 2016).
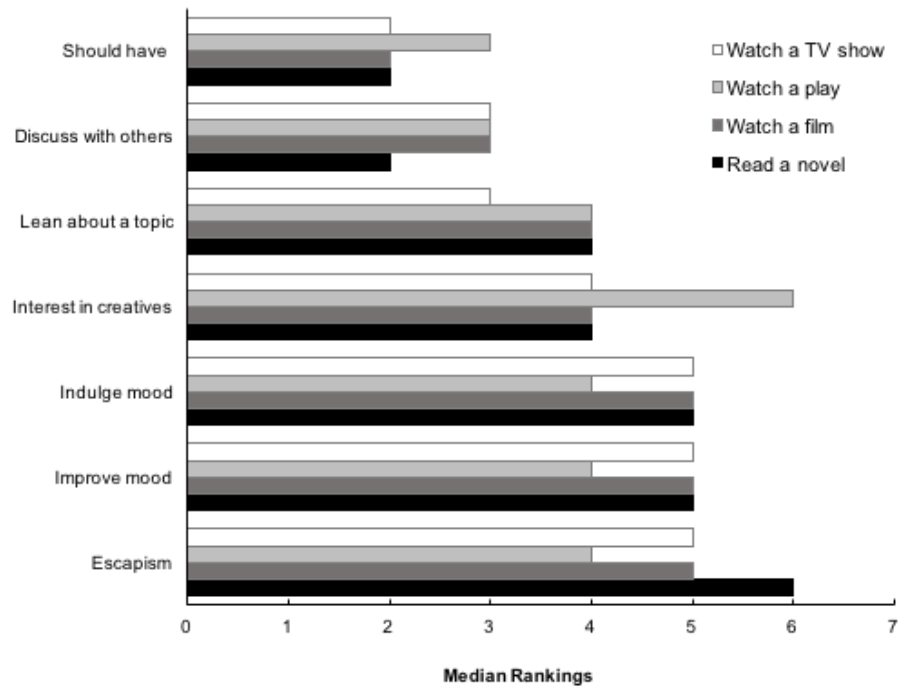
Figure 1. Median rankings for reasons participants tended to engage with four media formats (watch a TV show, watch a play, watch a film, read a novel).

Appendix C

SPSS syntax: Power analysis for Study 2.

Power analysis for group differences

```
* Encoding: UTF-8.
MATRIX DATA VARIABLES = Condition
ROWTYPE_ Score/FACTORS = Condition.
BEGIN DATA
1 N 6
2 N 6
1 MEAN 0.66
2 MEAN 0.25
. CORR 1
. STDEV 0.205
END DATA.
MANOVA Score BY Condition (1,2)
/MATRIX=IN(*)
/POWER exact t (.05)
/PRINT = Parameters (efsize) .
```

Power analysis for correlation

```
* Encoding: UTF-8.
INPUT PROGRAM.
LOOP n=100 TO 1000 by 50.
END CASE.
END LOOP.
END FILE.
END INPUT PROGRAM.
EXECUTE.
Compute r = 0.211.
Compute alpha = .05.
Compute tails = 2.
Compute r = abs (r) .
Compute t = (r* ( (n-2) **.5) ) / ( (1- (r**2) ) **.5) .
Compute d = 2*r / ( (1- (r**2) ) **.5) .
Compute delta = (d*sqrt (n-2) ) /2.
Compute alpha_tails = alpha/tails.
Compute fail = 1-alpha_tails.
Compute df = n-2.
Compute t_table = IDF.t(fail,df) .
Compute Power = 1-NCDF.t (t_table,df,delta) .
Execute.
Matrix.
GET M /VARIABLES=n d delta Power.
Print M/title = "Power Analysis for Correlation"
/clabels = "N" "d" "Delta" "Power"/format f10.4.
End Matrix.
```

Appendix D

Adaptation of Art-R used in Studies 2 and 3.The ART-R was published in *Research in Personality*, *40,* Mar, R. A., Oatley, K., Hirsh, J., dela Paz & Peterson, J. B., "Bookworms versus nerds: Exposure to fiction versus non-fiction, divergent associations with social ability, and the simulation of fictional social worlds", 694-671. Copyright Elsevier (2006). Adapted with permission from Elsevier.

Fiction

Romance	Sci-Fi/Fantasy	Suspense/Thriller	Domestic fiction	Foreign (translation)
Sidney Sheldon	Robert Jordan	Dean Koontz	John Updike	José Saramago
Danielle Steel	Douglas Adams	John LeCarré	W. O. Mitchell	Yukio Mishima
Jackie Collins	Anne McCaffrey	Robert Ludlum	Alice Munro	Gabriel García Márquez
Judith Krantz	William Gibson	Clive Cussler	Maeve Binchy	Albert Camus
Nora Roberts	Terry Brooks	Sue Grafton	Carol Shields	Umberto Eco
Iris Johansen	Terry Goodkind	Ian Rankin	John Irving	Milan Kundera
Diana Palmer	Piers Anthony	P. D. James	Toni Morrison	Paulo Coelho
Catherine Anderson	Arthur C. Clarke	John Saul	Amy Tan	W. G. Sebald
Joy Fielding	Ray Bradbury	Patricia Cornwell	Rohinton Mistry	Italio Calvino
Nicholas Sparks	Ursula K. Le Guin	Ken Follett	Sinclair Ross	Thomas Mann
E. L. James*	Kim Stanley Robinson*	Paula Hawkins*	Jodi Picoult*	Haruki Murakami*

Nonfiction Science	**Philosophy/Psychology**	**Political/Social-commentary**	**Self-help**	**Business**
Stephen Hawking	Roland Barthes	Noam Chomsky	Jack Canfield	Faith Popcorn
Stephen J. Gould	John Searle	Mary Beard*	Philip C. McGraw	Jim Collins
Richard Dawkins	Jean Baudrillard	Michael Moore	M. Scott Peck	Napoleon Hill
Thomas Kuhn	Michel Foucault	Eric Schlosser	Robert Fulghum	Robert T. Kiyosaki
Ernst Mayr	Bertrand Russell	Bob Woodward	Emma Bombeck	Stephen C. Lundin
Douglas Rushkoff	Antonio Damasio	Pierre Berton	Jean Vanier	Peter S. Pande
Amir D. Aczel	Daniel Goleman	Naomi Klein	Stephen R. Covey	Kenneth H. Blanchard
Matt Ridley	Jeffrey Gray	Naomi Wolf	Melody Beattie	Peter F. Drucker
John Maynard Smith	Joseph LeDoux	Robert D. Kaplan	Deepak Chopra	Barry Z. Posner
Diane Ackerman	Oliver Sacks	Susan Sontag	Marianne Williamson	Spencer Johnson**
Yuval Noel Harari*	Sam Harris*	Cordelia Fine*	Sarah Knight*	Sheryl Sandberg*

Foils				
Lauren Adamson	John Coundry	Martin Ford	James Morgan	Dale Blyth
Eric Amsel	Edward Cornell	Harold Gardin	Scott Paris	Robert Emery
Margarita Azmitia	Carl Corter	Frank Gresham	Richard Passman	Franklin Manis
Oscar Barbarin	Diane Cuneo	Robert Inness	David Perry	Alister Younger
Reuben Baron	Denise Daniels	Frank Keil	Miriam Sexton	Hilda Borko
Gary Beauchamp	Geraldine Dawson	Reed Larson	K. Warner Schaie	Frances Fincham
Thomas Bever	Aimee Dorr	Lynn Liben	Robert Siegler	Morton Mendelson
Elliot Blass	W. Patrick Dickson	Hugh Lytton	Mark Strauss	Steve Yussen

Note. *New name; **Amended.

Updates to ART-R (Mar et al., 2006). "M. D. Johnson Spencer" was amended to "Spencer Johnson". New names (10%) were added to each dimension (one in each genre). These were authors whose works have been published or re-published within 5 years of the time of scale construction (i.e., since 2013). Each critical dimension consisted of 55 names and there were 40 foils in total. "Norman Mailer" represented a nonfiction author on the original task; however, as he is a recipient of the Pulitzer Prize for both fiction and nonfiction and renowned for his use of literary devices in nonfiction writing, selections for this author could represent familiarity with either fiction or nonfiction, and so the name was replaced with a nonfiction author of the same genre ("Mary Beard").

Appendix E.1

Mean Absolute Difference Scores Presented for Each Level of Experiencer and Valence

within Each Strategy Group.

		Grand Mean	Simulation (i)	Theorisation (j)	Mean difference (i–j)	*p*
Experiencer	A	1.39	1.49	1.29	0.20 [0.07, 0.32]	.005
	B	1.74	1.80	1.67	0.13 [0.04, 0.23]	.01
	C	1.39	1.43	1.34	0.09 [-0.008, 0.18]	.06
	D	1.71	1.40	2.05	-0.65 [-0.78, -0.52]	.001
	E	1.15	1.31	0.98	0.33 [0.23, 0.43]	.001
	F	1.33	1.11	1.56	-0.45 [-0.55, -0.33]	.001
Valence	Positive	1.61	1.63	1.59	0.04 [-0.06, 0.16]	.40
	Neutral	0.89	0.79	0.99	-0.20 [-0.29, -0.11]	.001
	Negative	1.85	1.84	1.86	-0.02 [-0.14, 0.11]	.73
Experiencer × Valence	A × positive	1.90	1.73	2.08	-0.35 [-0.58, -0.14]	.007
	A × neutral	0.92	1.03	0.80	0.23 [0.08, 0.38]	.002
	A × negative	1.51	1.85	1.16	0.70 [0.51, 0.89]	.001
	B × positive	1.90	2.09	1.69	0.39 [0.24, 0.54]	.001
	B × neutral	1.12	1.03	1.21	-0.17 [-0.32, -0.03]	.03
	B × negative	2.13	2.17	2.09	0.09 [-0.12, 0.30]	.41
	C × positive	1.12	1.34	0.90	0.44 [0.31, 0.58]	.001
	C × neutral	0.91	0.62	1.22	-0.60 [-0.74, -0.47]	.001
	C × negative	1.95	2.12	1.77	0.35 [0.19, 0.51]	.001
	D × positive	2.06	1.58	2.56	-0.98 [-1.20, -0.76]	.001
	D × neutral	0.88	0.82	0.94	-0.12 [-0.26, 0.18]	.10
	D × negative	2.48	1.97	3.01	-1.04 [-1.34, -0.74]	.001
	E × positive	1.46	1.69	1.23	0.47 [0.33, 0.61]	.001
	E × neutral	0.66	0.67	0.65	0.02 [-0.12, 0.17]	.78
	E × negative	1.23	1.45	0.99	0.46 [0.32, 0.60]	.001
	F × positive	1.19	1.20	1.17	0.04 [-0.12, 0.20]	.64
	F × neutral	0.86	0.51	1.23	-0.72 [-0.86, -0.58]	.001
	F × negative	1.79	1.50	2.10	-0.60 [-0.78, -0.41]	.001

Note. 95% bias-corrected and accelerated confidence intervals (using $N = 1000$ bootstrapping) for mean differences are presented in brackets. Experiencer D rated each negative image at -4 (the most extreme negative rating possible). Therefore, participants' estimates could never be over-negative and so values or D × Negative could only be positive (and therefore identical to the absolute values).

Appendix E.2

Mean Raw Difference Scores Presented for Each Level of Experiencer and Valence within

Each Strategy Group.

		Grand Mean	Simulation (i)	Theorisation (j)	Mean difference (i–j)	*p*
Experiencer	A	-0.17	0.10	-0.46	0.56 [0.38, 0.74]	.001
	B	-0.19	-0.001	-0.39	0.39 [0.24, 0.51]	.001
	C	0.09	0.17	0.015	0.16 [-0.01, 0.32]	.46
	D	0.11	0.30	-0.10	0.41 [0.25, 0.56]	.001
	E	-0.20	-0.14	-0.27	0.13 [-0.01, 0.26]	.69
	F	0.40	0.26	0.55	-0.29 [-0.44, 0.16]	.001
Valence	Positive	-1.06	-0.76	-1.37	0.60 [0.44, 0.76]	.001
	Neutral	-0.10	0.21	-0.44	0.65 [0.51, 0.78]	.001
	Negative	1.18	0.90	1.48	-0.58 [-0.78, -0.38]	.001
Experiencer	A × positive	-1.67	-1.33	-2.03	0.69 [0.43, 0.97]	.001
× Valence	A × neutral	0.13	0.36	-0.13	0.49 [0.28, 0.70]	.001
	A × negative	0.92	1.16	0.66	0.51 [0.20, 0.79]	.003
	B × positive	-1.46	-1.35	-1.57	0.22 [-0.002, 0.41]	.055
	B × neutral	-0.21	0.30	-0.76	1.05 [0.84, 1.28]	.001
	B × negative	1.53	1.50	1.56	-0.06 [-0.31, 0.19]	.65
	C × positive	0.02	0.34	-0.31	0.64 [0.44, 0.85]	.001
	C × neutral	-0.52	-0.09	-0.98	0.89 [0.67, 1.10]	.001
	C × negative	0.61	0.24	1.00	-0.76 [-1.06, -0.49]	.001
	D × positive	-1.84	-1.17	-2.54	1.37 [1.14, 1.60]	.001
	D × neutral	-0.21	0.16	-0.60	0.76 [0.57, 0.94]	.001
	D × negative	2.48	1.97	3.01	-1.04 [-1.35, -0.70]	.001
	E × positive	-0.66	-0.47	-0.85	0.38 [0.15, 0.60]	.006
	E × neutral	0.02	0.34	-0.32	0.66 [0.49, 0.82]	.001
	E × negative	0.18	-0.18	0.56	-0.74 [-0.98, -0.51]	.001
	F × positive	-0.75	-0.50	-1.00	0.50 [0.27, 0.71]	.001
	F × neutral	0.13	0.16	0.12	0.04 [-0.14, 0.22]	.67
	F × negative	1.46	0.90	2.04	-1.14 [-1.40, -0.88]	.001

Note. 95% bias-corrected and accelerated confidence intervals (using $N = 1000$ bootstrapping) for the mean difference (simulation-theorisation) values are presented in brackets.

Appendix F

The Wallet: Stimulus story created for Study 3.

It was one of those rough, blustery Monday mornings when the wind seemed to find its way all around collars and right through scarves. The snow, which was gleaming white on Sunday afternoon, was melting now, and the pavements were lined with piles of grey ice that looked like big thick slugs. I walked past a red bobble hat wedged onto a melting ball of ice whose pebble smile was slowly sliding into a frown. I pulled my own hat down to my eyebrows; the wind was making my eyes water.

I'd made an effort to look smart this morning as I had a meeting first thing with my boss, Sam. It was the end of my trial period in the new job, and there were two of us hoping to be chosen for one permanent role. There was a chance I'd end the day with no job to go back to, which made me nervous, and the cold wasn't helping. *I'm sick of winter*, I thought, as a paving slab tipped beneath me, spitting stagnant water over my already-soggy shoes. I felt wet seep through the stitching and into my socks. *Looks like it's going to be one of those days,* I groaned.

I thought about taking the pavement along the main road, but I was running late so I decided to go through the park as usual. I was already drenched, and practically freezing. *I hope they've fixed the office radiator,* I thought. I was pretty much alone in the park—everyone else was probably driving today, or staying at home, lucky them. A familiar red-faced jogger ran towards me, the grass crunching under each *bof bof bof* of his trainers. As he rushed past I nodded to him, as I always do. 'iya', he panted, dropping the 'h'. His expression was vacant with focus and exhaustion, and a few moments later his *bof bof boffing* faded out behind me.

I checked my watch. Twenty-past eight—the train was due in five minutes! The station was just on the other side of the park, so I would make it, if I hurried. At the end of the park was a little mud path that led straight up to the platform. Taking this route would shave off a good three or four minutes, as long as it wasn't too muddy, so I headed in that direction. I hadn't brought an umbrella with me—the wind would have flung it inside out anyway—so I had settled on a woolly hat, which was fine if it didn't rain too heavily. It was hardly the height of fashion, in fact I don't even remember buying it. Someone must've left it at my house sometime, but a warm, dry head was more important than looking good in this weather.

I was just wondering who the hell this hat belonged to, when something ice-cold slithered down the back of my neck. I put my hand to it. Just drops of water. Ice *cold* water. 'Eurgh!' I murmured. I pictured one of those cartoon rainclouds hovering above my head and following me wherever I went. When I looked up, there was no raincloud, but a low-hanging branch of the huge old tree that marked the edge of the park. Its branches were twisted and brown and wet and its leaves quivered beneath melting ice. A huge, shabby blackbird screeched and flapped clumsily, and the branch quivered, icy water drenching me some more. 'Cheers for that!' I yelled. 'SQUAWK' it responded, as it flapped away into the grey air. My hat was soaked now, and no longer any use to me, so I took it off, wrung it out a bit and shoved it in my pocket. The squeaky texture of the wet wool set my teeth on edge.

I looked again at my watch. Only one minute until the train. The shortcut that lead to the platform was just ahead, and so with a sharp, cold intake of breath, I ran towards it. The mud had formed into a series of icy foot-shaped peaks and folds. Placing my feet in the imprints provided, and swinging my arms for momentum, I propelled myself up the slope. The rattle of the tracks signalled the approaching train. My foot slipped, I wobbled, waving my arms like a scarecrow to regain my balance. Just a few more steps—quick! Metal

screeched loudly against metal. I jumped off the frozen mud and onto the concrete platform. The train jolted to a halt and its doors opened: *Bing! Bing! Bing!* I paused just long enough to exhale. I made it. I jogged across the platform, stepped up over the gap and onto the train, jostling elbows with a man pushing his way on at the same time. 'Sorry' we mumbled at each other, without even looking up.

I leant against the wall by the door and breathed. *Not bad for someone who slept through their alarm*, I thought. It was a busy train; some kids in school uniform had made a seat of the luggage rack, and were climbing right in as if they were sharing a bunk-bed. A young-ish couple stood near me, next to the door. They were standing so close to each other that at first I thought they were kissing, until I heard 'just forget it,' and saw that they were arguing, quietly and in secret. Realising I was staring, I tried to move my gaze nonchalantly towards the doorway, as if it had merely, unseeingly, passed them by in its search for the view. Yellow lines, concrete, a puddle. Something beside the puddle. My hat? I felt in my pocket. No, still there, still damp. I squinted. Small, square, dark—a wallet! Someone had dropped their wallet! I looked around for someone looking walletless. Nothing. What should I do? I didn't want to just leave it there, but I needed to get to my meeting. I hesitated. Someone would be worried. Perhaps I should hand it in at the ticket office. That would be the right thing to do. If I didn't get it, who knows what would happen to it… I jumped down off the train just as the doors *Bing Bing Bing-ed* and closed behind me.

As the train engine groaned into action, I bent down to pick up the wallet. It was dry, so it couldn't have been there long. Already regretful—the next train was half an hour away—I walked towards the ticket office. The sign on the door read 'CLOSED'. *Huh.* I sat down on a cold, painted metal bench that sent a shiver through my legs, and opened up the wallet. It was nice, good quality leather or a very convincing fake. It had that musty leather smell, so I reckoned it was real. It was thick too, not with change, but with… cards? Yes, some. What else... notes! It was thick with notes! I counted them quickly, roughly*, ten twenty, forty, ninety…* there must have been two-hundred and fifty quid! Perhaps I could simply pocket some cash, and then leave it on the bench to be found…. No, I couldn't do that. I didn't know why someone had that much cash on them, but I couldn't be certain that they were wealthy enough to lose it. Perhaps they needed it to pay off a debt, or to buy a present for their kid. No, I had to return it.

I thumbed through the cards. Coffee shop card, supermarket points card, another coffee shop card, a passport sized photo of a baby… aha!… a driving licence. Bingo. I read the name: *Field, Mr James, R.* He looked smart in his photo and fairly young. Address: 13b Eric Street. Postcode: across town. I couldn't get across town, I *had to* get on the next train. I checked the inside pockets. More coffee cards, credit card, debit card, library card… and then a little pile of rectangular business cards, each one with two rounded corners, top right and bottom left, and each one belonging to *James Field, Armstrong & Grey Ltd., 44 High Street.* The high street was only five or six turnings along the main road, so I reckoned James Field dropped his wallet as he got off the train and he'd probably be able to come straight back to collect it. I inspected the card more closely, turning it between my finger and thumb like a detective eyeing the evidence. The design was minimal, with a grey wavy line through the middle, and there was no job title, but there was a mobile number. I pulled my glove off with my teeth, took out my phone and started dialling.

I was just preparing what to say to James Field's voicemail when the ringing stopped.
'Hello' came a man's voice.
'Oh er hello, I um I've found your wallet. It had your business card in it so I…'
'My wallet?'
'Yes, that's right. I thought I should call straight away as it has so much cash in it.'
There was a pause, followed by, 'hang on, let me check… oh gosh I've lost it'.

'Yes.'

'And you've found it?'

'Yes at Oakley station.'

'Oh goodness me, I didn't even know it was gone! Thank you so much for calling!' He sounded relieved.

'You're welcome', I said, 'are you nearby? I'm still at the station. Or should I bring it to you after work?' I cringed; *why did I say that*?

'Oh good grief no!' He responded, politely. 'I'm not far, I'll co me to you. Can you wait? I'll hop in a cab, so I'll be there a-sap.

'Yes, I can wait, my train's not till five-to,' I said.

'Great. See you soon. Thanks.' He hung up.

A door creaked behind me and the ticket office opened.

'Sorry', said a woman in train staff uniform, 'you haven't been waiting long have you? I was running late'. She held the door, and I walked towards the warm waiting area. 'Radiator's on, so you can get toasty,' she said.

I sat on another painted metal bench, but this one was warm. I stretched my legs out and felt the heat from the radiator on my feet. I rested for a few minutes, leaning back and soaking up the warmth as if I was on a beach. My ankles scorched, and I liked it. With a jolt of anxiety, I thought of the meeting I was about to be late to. I quickly typed a text message with my gloveless hand, and sent it to Sam.

RUNNING LATE, GOOD REASON, SEE YOU SOON FOR OUR MEETING. SORRY.

I felt nervous. As I tucked my phone back into my pocket, a man appeared in the doorway and looked at me. I stood up, holding the wallet in my other, gloved hand, and waved it at him.

'Oh you superstar!' the man said, gesturing at the wallet. It was James Field — I recognised him from his picture.

I smiled. 'It's no problem'.

James Field took the wallet, opened it and glanced inside the section containing all the cash.

'It's all there?' I asked. It sounded like a question, but I meant it as a statement: *I didn't steal any, if that's what you're thinking.*

'Yes —good,' said James Field. He paused for a moment, and gestured towards the platform, 'I hope I didn't make you late for anything'.

'Oh no… well… I missed my… but it's fine, I can get the next one…' I muttered.

'Oh goodness me, no! You must take my cab. It's just outside. I can call another.'

'Oh, thanks', I said. 'That would be a great help. It's just that I've got this meeting with my boss, and it's the end of my trial period and I'm a bit worried. You see, there are two of us going for the same job and I really want to be the one that gets it… and now I'm running late and my boss *really* hates lateness… and so I was um just a bit um worried that I would be late. So um… yeah, thanks!' I looked at the floor. *Shut up, idiot,* I thought.

After a moment's pause, James Field held out his hand to me. I shook it, instinctively. 'I'm James. Thank you so much for looking after my wallet for me,' he said, slowly and with kindness. 'I have a cab outside and I would be so happy if you would let it take you to work. On me, of course. And as an extra thank you, because you looked after my wallet for me even though it made you late for your meeting, I would like to give you this.' I looked up then at James's outstretched hand. Folded between two fingers was a little wad of notes.

'Oh no!' I protested, shocked. 'The cab's great, though.'

'Please', said James, seeming like he really meant it. 'I know it's a rather tacky way to say thank you, but given that you're in a rush I'd really appreciate it if you would take it. Your honesty and kindness has saved me so much worry and hassle, and I really *want* to say thank you. I've got the money, so please take it.'

I smiled and reached forward. 'Ok… thank you,' I said.

'No, thank *you!'* Said James, and ushered me towards the taxi.

A few minutes later I was rolling along the main road towards work. The car was warm, I could feel my hands again, and the driver had the radio on: '…we can expect clearer skies by the afternoon and even some sunshine in parts'. I smiled, looking out at the melting snow. I checked my watch—I was going to be on time after all, just as if I'd caught the train. *I'd better text Sam again*, I thought.

HI SAM, ME AGAIN. PLEASE IGNORE MY LAST MESSAGE — I WILL BE THERE ON TIME FOR OUR MEETING AFTER ALL.

As I put my phone away, I remembered that I was still holding the cash. I'd been so taken aback by James Field's generosity, that I hadn't even thought to count it! *Ten, twenty, seventy, one-hundred*! I gasped. I felt like I'd won the lottery! *Well at least if I get fired, I'll have this to tide me over!* I thought. *Wow. James Field. What a decent bloke.*

'You alright?' Asked the taxi driver—he must have heard me gasp.

'Yeah, I am.' I replied, grinning, and meaning it.

He smiled back, and pulled around the corner, coming to a halt right outside my office building. 'Here you are', he said.

As I stepped out of the cab my phone buzzed; it was Sam, replying to my text. I felt my stomach tense. What if this had tipped it, and I was the one to be fired? Perhaps she'd tell me not to bother coming back at all. I opened the message.

DON'T WORRY, I'M STUCK IN BAD WEATHER AND RUNNING 5 MINS LATE. THANKS FOR FINDING A WAY TO GET HERE ON TIME, YOU'LL BE GETTING MY JOB NEXT IF I'M NOT CAREFUL!

My stomach relaxed. The driver pulled into the car park and up to the main entrance. 'It's paid for. Have a nice day!'

Stepping out, I waved back. 'I will,' I replied. 'You too.' I breathed deeply, and thought what an eventful day it had been already, and it wasn't even nine o' clock! *It's going to be one of those days,* I thought, *one of those very good days.*