

**Cross-Cultural Investigation Into the Associations of Fiction Reading Habits With Mentalizing Skills
and Stereotyping Among Adults in the United Kingdom and Japan**

Atsunobu Suzuki¹, Hidekazu Osanai², and Chang Hong Liu³

¹ The University of Tokyo, Japan

² Kio University, Japan

³ Bournemouth University, United Kingdom

Author Note

This work was supported by the Japan Society for the Promotion of Science (KAKENHI grant numbers 21K03124, 21H05323, 23H03878) and the University of Tokyo Humanities Center (LIXIL Ushioda East Asian Humanities Initiative Open Research (A) (October, 2019–September, 2020)). We thank Becky Gilbert for providing the English version of the Author Recognition Test and Makiko Yamada for permission to use the Japanese version of the “Reading the Mind in the Eyes” test.

Correspondence concerning this article should be addressed to Atsunobu Suzuki, Department of Psychology, Graduate School of Humanities and Sociology, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. Email: atsuzuki@l.u-tokyo.ac.jp

Abstract

Fiction reading habits are thought to be associated with a favorable social-cognitive profile, including increased mentalizing skills and decreased stereotypical beliefs. However, the available evidence for this association is largely based on a specific task, the “Reading the Mind in the Eyes” Test (RMET), and on data collected from Western populations. This raises questions regarding the generalizability of the findings. We addressed this issue by examining the relationships between fiction reading habits (as measured by the Author Recognition Test) and several social cognitive variables, such as basic emotion recognition from facial expressions and the acceptance of stereotyping, in adults in the United Kingdom (UK) and Japan. Among UK adults, fiction reading habits were positively correlated with performance on both the RMET and the basic emotion recognition task and negatively correlated with the acceptance of stereotyping in general, although their relationship with facial stereotyping (common beliefs linking facial appearance to personality) was unclear. Meanwhile, these relationships were not statistically significant in Japanese adults. Thus, in the UK, the positive association of fiction reading habits with mentalizing skills seems to generalize beyond the RMET, whereas the negative association with stereotypical beliefs may not hold for facial stereotyping. The lack of similar associations in Japan may reflect differences in the measurement materials and/or storytelling traditions between the two countries, highlighting the importance of further research in non-Western populations.

Keywords: cross-cultural study, fiction reading habits, mentalizing, social cognition, stereotyping

1 It has been argued that reading fiction, one of the most popular forms of leisure, improves the
2 mental processes and skills involved in understanding and interacting with others (social cognition),
3 such as inferring the mental states of others (mentalizing/theory of mind) and feeling and thinking as
4 others do (empathy) (Mar, 2018; Mar & Oatley, 2008). Such effects of fiction are proposed to follow
5 the “use it and improve it” principle and have been compared to those of flight simulation in pilot skills
6 training (Oatley & Djikic, 2018). As fiction is typically about people and their mental states and rela-
7 tionships, reading fiction naturally engages the neurocognitive processes involved in social cognition
8 (Lehne et al., 2015; Tamir et al., 2016), potentially boosting these processes. Further, it has been sug-
9 gested that literary (e.g., award-winning) fiction is more effective in promoting social cognition than
10 popular (e.g., best-selling) fiction (Castano, 2024). This proposal is based on the idea that literary fic-
11 tion has more features than popular fiction that elicit extensive mental inferences in readers, such as
12 ambiguous and complex characters that resist simple schematic understanding (Kidd & Castano, 2013,
13 2017; Koopman & Hakemulder, 2015) and metaphorical implicit descriptions of characters’ minds us-
14 ing concrete language (Castano, 2024).

15 There is empirical evidence of positive social cognitive outcomes associated with reading fic-
16 tion (Eekhof et al., 2022; Mar, 2018). Studies have shown that fiction reading habits are positively
17 correlated with self-reported empathy and performance in a mentalizing task (Mumper & Gerrig,
18 2017). This association is maintained even when the participants’ age, gender, personality, and general
19 intelligence are statistically controlled (Fong et al., 2013; Kidd & Castano, 2017; Mar et al., 2006, 2009).

Experimental studies have shown causal increases in empathy and mentalizing performance after reading fiction (Dodell-Feder & Tamir, 2018). Even reading a short passage of fiction has been reported to improve performance on certain mentalizing tasks (Kidd & Castano, 2013). However, such brief exposure effects have often not been replicated in subsequent studies (Panero et al., 2016; Samur et al., 2018), and meta-analytic reviews have suggested that these effects are small (Dodell-Feder & Tamir, 2018) and that the relevant evidence is not robust (Quinlan et al., 2023). Moreover, empirical evidence is inconclusive regarding the differential effects of reading literary versus popular fiction on mentalizing and empathy (Dodell-Feder & Tamir, 2018; Fong et al., 2013; Kidd & Castano, 2013; Mumper & Gerrig, 2017), and it remains unclear how best to distinguish between these two types of fiction (Koopman & Hakemulder, 2015; Mar, 2018).

Researchers are now shifting their focus to identifying the moderators of the associations between fiction reading and social cognition (Eekhof et al., 2022; Mar, 2018). Meta-analyses have shown that the aggregate effect sizes for the associations are statistically significant but small (Mumper & Gerrig, 2017 [correlations around or less than .2]; Dodell-Feder & Tamir, 2018 [Hedges' g s around or less than .2]). However, the small aggregate effects can be strengthened (or further weakened) by the characteristics of the text, the reader, and the social cognitive task (Eekhof et al., 2022; Mar, 2018). For example, previous research suggests that the relationship between reading fiction and social cognition may vary according to the genre of fiction (Black et al., 2018; Buttrick et al., 2023; Kidd & Castano, 2013; Fong et al., 2013), transportation (Lenhart et al., 2022) and eudaimonic experiences (De Mulder

et al., 2022) during reading, the reader's openness to experience (Djikic et al., 2013), and the types of measures of social cognition (Stansfield & Bunce, 2014). Thus, investigating the moderators and boundary conditions of the positive influence of fiction on social cognition is increasingly being viewed as an important line of research in the field.

An often-cited limitation of previous studies is their heavy reliance on a single task—the “Reading the Mind in the Eyes” Test (RMET; Baron-Cohen et al., 2001)—to obtain a performance-based measure of social cognition (Eekhof et al., 2018; Mar, 2018). In this test, participants view photographs of different sets of human eyes and choose a word describing what each of these individuals is thinking or feeling out of four response choices. The RMET is designed as a test of theory of mind or mentalizing, that is, the ability to infer the mental states of others (Baron-Cohen et al., 2001; Dodell-Feder et al., 2020), although it may primarily assess the inference of affective states (i.e., emotion recognition; Castano, 2024; Oakley et al., 2016). Performance-based measures of social cognition have an important advantage over self-report measures, which are more susceptible to contamination by socially desirable responses and inaccurate self-concepts. Nevertheless, performance on the RMET has been shown to have positive correlations with measures of verbal ability (Peterson & Miller, 2012). This association is logical given that the RMET uses complex and unusual vocabulary as descriptors of mental states (e.g., “aghast”; Dodell-Feder et al., 2020). Consequently, the well-replicated positive association between reading fiction and performance on the RMET (Mumper & Gerrig, 2017) may reflect improved verbal skills through reading rather than improved social cognition (Eekhof et al., 2022).

1 The first goal of this study was to examine the relationships between fiction reading habits and
2 performance on another mentalizing task that tests the recognition of basic emotions from facial ex-
3 pressions (Ekman, 1992). Basic emotion terms (e.g., “angry,” “happy”) are more common than the
4 mental state descriptors used in the RMET (Dodell-Feder et al., 2020). In addition, scores on basic
5 emotion recognition tasks are less related to age and education than those on the RMET, suggesting
6 less reliance on verbal ability (Dodell-Feder et al., 2020). Two experiments conducted by Schwering et
7 al. (2021) yielded conflicting results regarding the correlation between fiction reading habits and the
8 recognition of “simple” emotions (anger, joy, surprise, disgust, sadness, fear) from audiovisual stimuli;
9 the correlation was insignificant in the first experiment, but positive in the second experiment. De
10 Mulder et al. (2022) reported that among adolescents, mentalizing skills (including basic emotion
11 recognition) were not positively associated with overall fiction reading frequency. However, these skills
12 were positively associated with the frequency of eudaimonic experiences (e.g., deeper understanding
13 of the meaning of life) and hedonic experiences (e.g., feelings of pleasure and excitement) while read-
14 ing. Given these mixed and nuanced findings, further research examining the relationship between
15 fiction reading habits and basic emotion recognition from facial expressions would be a valuable addi-
16 tion to the literature.

17 Another goal of this study was to test whether reading fiction is associated with reduced
18 stereotyping, which is an association of growing interest (Eekhof et al., 2022). Theoretically, ste-
19 reotypical perceptions of people (overgeneralized views of others) may be mitigated by reading

1 fiction in at least two ways. First, it has been argued that fictionality allows people to empathize
2 with others without worrying about real-world consequences such as misinterpretation and
3 moral obligation (De Mulder et al., 2022; Koopman & Hakemulder, 2015). It has also been pro-
4 posed that facilitating empathy with out-group others reduces one's stereotyping of them (Batson
5 & Ahmad, 2009; Mar & Oatley, 2008). Indeed, Johnson et al. (2013) showed that exposure to a
6 passage from a novel depicting a counter-stereotypical Muslim woman led to increased empathy
7 for Arab Muslims and decreased stereotyping of them among predominantly White participants
8 (see also Vezzali et al., 2015). Second, because fiction typically depicts the complex and nuanced
9 nature of human beings (Mar, 2018), prolonged exposure may teach readers about the limitations
10 of social categories in understanding people and discourage the use of stereotypes. Such effects
11 may be more pronounced for literary fiction than for popular fiction, as characters in literary fic-
12 tion tend to be more "category-resistant" (less easily fitted into social categories) than those in
13 popular fiction (Castano et al., 2021). Evidence for this prediction comes primarily from studies
14 examining how fiction reading habits correlate with self-report measures of stereotyping and
15 other simplistic and fixed views of people and the world. Specifically, fiction reading habits have
16 been shown to be correlated with higher gender role egalitarianism and lower gender role stere-
17 otyping (Fong et al., 2015), higher attributional complexity (a preference for complex explanations
18 of human behavior; Buttrick et al., 2023; Castano et al., 2020), lower psychological essentialism
19 (the belief that people fall into discrete categories; Buttrick et al., 2023; Castano et al., 2021), and

1 lower dysfunctional relationship beliefs (e.g., “disagreement is destructive”; Stern et al., 2019).

2 Some of these studies also showed stronger associations of literary fiction with reduced stereo-

3 typing than popular fiction (Buttrick et al., 2023; Castano et al., 2020, 2021), whereas others did

4 not (Fong et al., 2015; Stern et al., 2019). Given the growing interest in these associations (Eekhof

5 et al., 2022) and the vulnerability of self-reports to desirable responses and inaccurate metacog-

6 nition, a study using a performance-based measure of stereotyping to examine its relationships

7 with fiction reading would be fruitful.

8 To accomplish the above goal, this study focused on “facial stereotyping.” People hold consen-

9 sual beliefs that associate certain sets of facial features with certain psychological traits (Todorov et al.,

10 2015). For example, individuals with cheerful and feminine faces tend to be perceived as more trust-

11 worthy (Oosterhof & Todorov, 2008). Facial stereotyping is considered to have dubious validity and

12 little practical use for trait inference in the real world (Olivola et al., 2014). Suzuki et al. (2022) demon-

13 strated temporally stable individual differences in the tendency to make extreme trait judgments

14 based on facial stereotyping. This tendency, or face-based trait inference (FBTI) extremity, is derived

15 from performance on a face judgment task and has been shown to be immune to socially desirable

16 responses. As fiction typically teaches readers about the complexities of human psychology (Mar,

17 2018), it may reduce their reliance on facial stereotyping. However, depicting the facial appearance of

18 story characters is a classic technique that helps readers infer the kind of people they are (Kronshage,

19 2018; Tytler, 1982), which may reinforce facial stereotyping. Thus, the relationships between fiction

reading habits and the degree of acceptance of facial stereotyping are not obvious and worthy of empirical investigation.

Furthermore, relatively little research has been conducted on whether and how the relationships between fiction reading and positive social outcomes hold outside of Western countries. Of the 23 published articles included in the meta-analyses by Dodell-Feder and Tamir (2018) and Mumper and Gerrig (2017), 18 reported on studies conducted in Western Europe and North America. Although the nationalities of the participants in the remaining five studies are unclear, English materials were used to measure and/or manipulate fiction reading. A few studies have examined this issue in non-Western countries, using different methods. Tabullo et al. (2018) studied Argentine adults and found a significant positive association between fiction reading habits and performance on the RMET for male participants, but not for female participants. They speculated that men might be more susceptible to the effects of fiction reading than women because of their generally lower openness to experience. Takahashi et al. (2023) studied Japanese undergraduates and showed that RMET performance had only a weak or no correlation with fiction reading habits ($r = .11$ in Study 1 and $r < .01$ in Study 2) and did not improve after brief exposure to fiction. The results suggest that the effects of fiction reading may be truly null in the population studied or may have been masked by the low internal consistency of the RMET or some procedural factors (e.g., duration of fiction exposure). Kevane (2020) reported that in Burkina Faso, a multi-month program to provide access to books in rural villages without libraries resulted in increased reading among adolescents but did not promote prosocial behaviors

(trust and cooperation) that are thought to be related to empathy (Batson & Ahmad, 2009). The limited exposure of the target population to fiction during childhood and the quality of fiction books by local authors used in the program may explain the negative result. Overall, the relationships between fiction reading and social cognition appear to be less reliable in non-Western countries than in Western countries for various and unclear reasons, supporting the need for further cross-cultural research.

The final goal of this study was to determine whether the associations between fiction reading habits and positive social cognitive outcomes noted in Western countries could be generalized to an Eastern country. To achieve this goal, adults in the United Kingdom (UK) and Japan were studied. It has been argued that Western and Eastern cultures place greater emphasis on independence and interdependence, respectively, in social relationships (Markus & Kitayama, 2010). The independent view construes social relations as being primarily driven by personal preferences and goals, whereas the interdependent view focuses on the role of communal and relational concerns. It has been argued that the spectrum of independence and interdependence is an important dimension along which social cognition can vary (Fiebich, 2016; Lillard, 1999; Slaughter & Perez-Zapata, 2014). Fiebich (2016) argues that the independent view leads to mentalistic narratives that explain human behavior based on mental states, whereas the interdependent view favors socio-situational narratives that emphasize social roles and norms. For example, Western children often refer to the mental states of characters when explaining their responses to false belief tasks (Bartsch & Wellman, 1989), whereas Japanese children often provide social rule-based justifications such as “they promised” (Naito & Koyama, 2006). Thus, a

comparison between Western (e.g., the UK) and Eastern (e.g., Japan) countries could serve as a valuable minimal test of the universality of the relationships between fiction reading habits and social cognition (Majid, 2023).

In summary, the present study examined the relationships between fiction reading habits and various social cognitive variables among adults in the UK and Japan to answer the following three research questions. Question 1: Are fiction reading habits positively related to performance on a mentalizing task beyond the RMET (i.e., recognizing basic emotions from facial expressions)? Question 2: How are fiction reading habits related to (facial) stereotyping? Question 3: Are the associations between fiction reading habits and positive social cognitive outcomes observed in Western countries replicable in an East Asian country (Japan)? This paper reports on an exploratory study that was part of a larger project preregistered on the Open Science Framework (Suzuki et al., 2024).

Method

Participants

UK and Japanese participants were recruited through Prolific Academic (<https://www.prolific.co/>) and CrowdWorks (<https://crowdworks.jp/>), respectively. The UK participants were native English speakers living in the UK. The Japanese participants were native Japanese speakers living in Japan. They received approximately £4.4 (UK sample) or ¥600 (Japanese sample) as an honorarium for their participation in the study. Eligible participants for this study were 20 years of age or older. Data collection in each country was terminated when the number of participants reached 318, which took one

week in the UK and one day in Japan. Assuming an effective response rate of 90%, the sample size was determined to detect a correlation of $r = .211$ (i.e., the overall correlation between fiction reading habits and performance on the RMET reported in Mumper and Gerrig's [2017] meta-analysis) with 95% power at the 5% significance level.

After completing data collection, we applied the following exclusion criteria: (1) accessing the online survey with a smartphone or a mobile device of similar (smaller) size; (2) failing or refusing to complete the entire survey; (3) selecting the same response category (other than the midpoint category) for all items in any of the tasks except for those questionnaires without reverse items; (4) agreeing with all NO items and disagreeing with all YES items in the Physiognomic Belief Scale (details explained later); and (5) failing any one of the attention check questions. Criterion 1 was imposed because some of the tasks required the recognition of subtle differences between face images. Criterion 2 meant that we analyzed complete data with no missing values. Criteria 3–5 were to exclude inattentive participants. These preregistered criteria resulted in the exclusion of 18 UK and 15 Japanese participants. In addition, two other participants were excluded from the Japanese sample, one because of missing data due to a technical error and the other because of duplicate participation in the survey. The remaining participants included in our subsequent data analyses were 300 UK (male: 93, female: 206, other: 1; Asian: 12, Black: 7, Mixed: 11, White: 267, Another: 2, Prefer not to say: 1; years of age: 20–70 [$M \pm SD = 34.88 \pm 10.98$]; years of schooling: 6–20 [14.72 ± 3.24]) and 301 Japanese participants (male: 122, female: 179; years of age: 20–72 [39.42 ± 9.73]; years of schooling: 6–20 [14.86 ± 2.39]).

The UK participants were more likely to be female ($\chi^2(1) = 5.799, p = .016$) and younger ($t(599) = -5.358, p < .001$) than Japanese participants, while years of schooling were comparable between them ($t(599) = -0.575, p = .565$).

Procedure and Tasks

Overview

Both the UK and Japanese surveys were conducted using the Qualtrics software (<https://www.qualtrics.com>). Each survey began with an informed consent form and some basic—mostly demographic—questions. This was followed by the tasks and questionnaires measuring fiction reading habits, mentalizing skills, stereotyping, and response biases. The order of the tasks and questionnaires was randomized. The order of the stimuli/items in each task/questionnaire was also randomized, unless otherwise noted.

Fiction Reading Habits

Author Recognition Test (ART). The ART provides an indirect measure of fiction reading habits based on readers' familiarity with fiction authors' names (Stanovich & West, 1989). Such an indirect measure is important because fiction reading habits are socially desirable and people may exaggerate them in self-reports (Stanovich & West, 1989). The ART has been shown to be a valid tool for measuring reading volume (Moore & Gordon, 2015) and has been used to examine the associations between fiction reading habits and social cognition (Mumper & Gerrig, 2017).

In the ART administered in the current study, participants read a list of names of over 100

people, approximately half of whom were real authors and the rest of whom were not. They were asked to select the names of those whom they knew to be real authors while being discouraged from guessing. Performance on this test was scored as the proportion of correct selections of real author names minus the proportion of incorrect selections of non-author names, which could range from -1 to 1 . Versions of the ART adapted to the UK (Gilbert et al., 2018; 65 real authors and 65 non-authors; Supplementary Table 1) and Japanese (Osanai et al., 2015; 52 real authors and 52 non-authors; Supplementary Table 2) cultures were used.

Questionnaire on Reading Habits. A two-part self-report questionnaire was administered as a complementary measure of reading habits. In the first part, participants indicated the approximate number of books they typically read during their leisure time each year. Separate responses were collected for the following nine genres: classics, modern/contemporary literary fiction, romance, fantasy, science fiction, suspense/thriller/horror/mystery, other genres of popular fiction, nonfiction, and others (presented in this fixed order). The first six genres were based on those listed in the Genre Familiarity Test (Black et al., 2018; Stern et al., 2019) with minor modifications (changing contemporary fiction to modern/contemporary fiction and combining suspense/mystery/thriller and horror into one genre). The last three genres were added to make the list more comprehensive. For each book genre, participants chose a number from 0 to 99 or selected the “100 or more” option. In the second part, participants indicated the frequency of their engagement in the following four reading-related activities: visiting a bookstore, visiting a library, browsing an online bookseller, and checking book reviews

(presented in this fixed order). These activities were based on items used by one of the authors to quantify reading habits in Japan. Responses were recorded on the following five-point scale: “never” (0), “only occasionally” (1), “more than once a month” (2), “more than once a week” (3), and “every day” (4).

Mentalizing Skills

Reading the Mind in the Eyes Test (RMET). The English (Baron-Cohen et al., 2001) and Japanese (Hirao et al., 2008) versions of the revised RMET were used to measure the ability to infer the mental states of other people from their eyes. In each trial, a picture of a Caucasian person’s eyes was displayed along with a list of four words describing certain mental states (e.g., aghast, fantasizing, impatient, alarmed). Each participant chose the word that, according to them, best described what the person in the picture was thinking or feeling. A total of 36 pictures of eyes were presented in the same fixed order as specified in the test manual. The Japanese version of the ART used the same pictures as the English version, and the words describing mental states were translated from English to Japanese. The proportion of correct responses was calculated as a measure of mentalizing skills.

Basic Emotion Recognition Task. Recognition of basic emotions from facial expressions was tested to obtain another measure of mentalizing skills. The task was adapted from Suzuki et al. (2022), and previous studies have used similar tasks (e.g., Suzuki et al., 2006a; Young et al., 1997). In this task, participants were presented with photographs of facial expressions along with a list of five emotion words: sadness, anger, happiness, fear, and disgust. For each facial expression, the participants rated

1 how well each emotion matched the affective experience that the person in the picture might be feel-
2 ing (Supplementary Figure 1A). Ratings were made on a six-point scale ranging from “does not match
3 at all” (1) to “matches very well” (6). A total of 25 face images were presented as stimuli. Face images
4 of a Caucasian male, “JJ,” from the Pictures of Facial Affect (Ekman & Friesen, 1976) were used in the
5 UK study and those of a Japanese female, “KS” (Suzuki et al., 2006a), in the Japanese study. The mor-
6 phed facial expressions (see below) of JJ and KS have been shown to be effective in detecting subtle
7 impairments in emotion recognition among patients with brain damage in the UK (Calder et al., 1996;
8 Sprengelmeyer et al., 1996) and Japan (Suzuki et al., 2006b). Suzuki et al. (2010) also showed that the
9 morphed facial expressions of both the models were effective to quantify individual differences in
10 emotion recognition among Japanese adults. Therefore, we believed that using the morphed facial
11 expressions of JJ and KS would be useful for our study, despite the gender differences. The stimuli
12 included one image of a prototypical facial expression (according to Paul Ekman’s theory; Ekman et al.,
13 2002) for each of the five emotions listed above. The remaining 20 photos were ambiguous facial ex-
14 pressions created by mixing the two prototypical expressions using a morphing software
15 (<https://www.fantamorph.com>). Specifically, for each of the 10 possible pairs of the five prototypical
16 facial expressions (e.g., sadness and anger), two intermediate expressions were created by mixing the
17 paired expressions in ratios of 3:2 (e.g., 60% sadness and 40% anger) and 2:3 (e.g., 40% sadness and
18 60% anger). Research has shown that in real life, people express a combination of different emotions
19 on their faces, including a mixture of positive and negative emotions, such as a happy and

simultaneously disgusted facial expression (Martinez, 2017). The ambiguous facial expressions created by our morphing protocol can be considered approximations of such compound facial expressions. They have been shown to be more useful for measuring emotion recognition performance than prototypical facial expressions, as performance often shows a ceiling effect for the latter but not for the former (Suzuki et al., 2006a). Furthermore, some clinical conditions associated with social affective problems are known to impair emotion recognition from morphed ambiguous facial expressions (Calder et al., 1996; Sprengelmeyer et al., 1996; Suzuki et al., 2006b), supporting the interpretation that our task measures some kind of emotional ability.

Recognition performance was first quantified for each emotion. Taking “anger” as an example, anger ratings were averaged separately for the nine “anger-expressing faces” (i.e., one prototypical angry facial expression and eight mixed facial expressions generated from it) and for the remaining 16 “anger-free faces.” The difference between the mean anger ratings for the anger-expressing and anger-free faces was then used as a measure of anger recognition ability. Other emotion recognition scores were calculated similarly. These emotion-specific scores were then averaged to obtain an overall index of facial emotion recognition ability, which could range from –5 to 5. Higher scores on this test indicate a higher level of ability to discriminate when certain emotions are expressed in facial expressions and when they are not, thus indicating better emotion recognition performance.

Stereotyping

Facial stereotyping was measured using an FBTI task and self-report questionnaires. A general

measure of stereotyping was also obtained using a self-report questionnaire.

Face-Based Trait Inference (FBTI) Task. This task was adapted from Suzuki et al. (2019, 2022) to assess the FBTI extremity (a task-based measure of facial stereotyping). In each trial, two faces were presented side by side on the screen (Supplementary Figure 1B). Participants compared their impressions of the two faces on a particular trait (e.g., trustworthiness) and indicated the extent to which they thought one person in the pair was more likely to have that trait than the other on a six-point scale (e.g., a scale ranging from “the person on the left is far more trustworthy” to “the person on the right is far more trustworthy”, with no neutral option such as “both people are equally trustworthy”). The traits assessed in this task were trustworthiness, dominance, and competence, which constitute a minimal set of universal dimensions of social cognition (Sutherland et al., 2016). The task was divided into three blocks, each of which required participants to rate six face pairs (three male and three female) on a single trait. The order of the blocks (i.e., the traits to be rated) was randomized. Each face pair consisted of two faces that gave opposite impressions on a relevant trait (Supplementary Figure 1B).

Photographs of Caucasian and Japanese faces were used as stimuli in the UK and Japanese studies, respectively. Caucasian faces were selected from the database created by Lundqvist et al. (1998), and Japanese faces were selected from an in-house database developed by the first author. The faces in these databases had been rated for impressions of trustworthiness, dominance, and competence, and the stimulus selection was performed using the existing rating data.

Each participant's FBTI extremity was scored using the following procedure (Suzuki et al., 2019, 2022). First, the ratings were recoded such that higher values reflected a greater tendency to make extreme and stereotype-consistent FBTIs. For example, when the more trustworthy-looking face of the face pair was shown on the left side of the display, the choice "the person on the left is far more trustworthy" was scored as six (i.e., the maximum score), whereas the choice "the person on the right is far more trustworthy" was scored as one (i.e., the minimum score). Second, the recoded ratings were averaged separately for each trait dimension, resulting in three trait-specific extremity scores. Finally, these trait-specific scores were averaged to produce a single measure of the FBTI extremity, which could range from 1 to 6. Higher scores (>3.5) on this measure indicate a strong tendency to make stereotypical inferences (e.g., judging trustworthy-looking faces as trustworthy), lower scores (<3.5) indicate a strong tendency to make anti-stereotypical inferences (e.g., judging trustworthy-looking faces as untrustworthy), and scores near the midpoint (3.5) indicate an indecisive or agnostic attitude toward FBTI. Previous studies (Suzuki et al., 2019, 2022) have shown that similarly measured FBTI extremity is related to physiognomic belief (the belief that various traits can be inferred from faces; see below for details), perceptual reliance (the reliance on perceptual cues in general when making social judgments; Livingston, 2001), and the tendency to make social judgments from social categories such as gender. These can be taken as support for our interpretation that the FBTI extremity score represents a tendency to accept (facial) stereotyping.

Questionnaires on Physiognomic Belief. Two questionnaires were administered to obtain self-

reported measures of facial stereotyping, both of which assessed the strength of the so-called physiognomic belief that various traits can be inferred from faces. One questionnaire, the Physiognomic Belief Scale (Suzuki et al., 2019), contained 14 “target items” that stated that certain traits could be inferred from faces (e.g., “I know a trustworthy person when I see their face”). Participants rated how true each item was to their thoughts on a four-point scale ranging from disagree (1) to agree (4). The ratings were averaged across the target items for each participant, and the resulting mean served as an index of the strength of the physiognomic belief. The questionnaire also included four “control items,” consisting of two “YES items” that should normally be accepted (e.g., “I can tell babies and the elderly apart by their faces”) and two “NO items” that should normally be rejected (e.g., “I can guess a person’s address correctly when I see their face”). These items were used for screening purposes.

The other questionnaire, developed by Jaeger et al. (2020a), consisted of three general statements about the feasibility of FBTI (e.g., “I can learn something about a person’s personality just from looking at his or her face”); for the Japanese study, the questionnaire was translated to Japanese with a back-translation check. Participants were asked to imagine that they were looking at the passport photo of a stranger and to indicate how much they agreed with each statement on a seven-point scale ranging from “strongly disagree” (1) to “strongly agree” (7). The average of the ratings quantifies the strength of the physiognomic belief.

Acceptance of Stereotyping. The Acceptance of Stereotyping Questionnaire (ASQ; Carter et al., 2006) was used as a general measure of stereotyping. It contained 12 statements that

(dis)supported the subjective validity, usefulness, and necessity of stereotypes (e.g., “Stereotypes are useful in daily life even though they are not always correct”); for the Japanese study, the questionnaire was translated to Japanese with a back-translation check. Participants rated how true each item was to their thoughts on a five-point scale ranging from “strongly disagree” (1) to “strongly agree” (5). Ratings were averaged across items such that the resulting mean reflected each participant’s favorable attitude toward stereotypes.

Response Bias

Balanced Inventory of Desirable Responding. Social desirability bias (people’s tendency to answer questions in a way that makes them look good) may inflate the correlations between some variables of interest (e.g., fiction reading habits and stereotyping). To measure and control for this bias, the Impression Management subscale of the Balanced Inventory of Desirable Responding (Paulhus, 1991) was used. The questionnaire contained 8 (English version; Hart et al., 2015) or 12 (Japanese version; Tani, 2008) items describing either desirable (e.g., “I never cover up my mistakes”) or undesirable (e.g., “I have said something bad about a friend behind his or her back”) behaviors. Participants indicated how much they agreed with each item on a scale ranging from “not true” (1) to “very true” (7). The proportion of extreme responses (i.e., the sixth and seventh interval choices for desirable behaviors and the first and second interval choices for undesirable behaviors) was computed for each participant as a measure of social desirability bias.

Extreme Response Style Measure. Extreme response bias (people’s tendency to select

extreme rating categories, regardless of item content) may inflate correlations between some variables of interest (e.g., FBTI extremity and basic emotion recognition). To measure and control for this bias, the Extreme Response Style Measure (Greenleaf, 1992) was used; for the Japanese study, the questionnaire was translated to Japanese with a back-translation check. The questionnaire consisted of 16 unrelated items (e.g., “When I see a full ashtray or waste-basket, I want it emptied,” “I am a homebody”). Participants indicated how strongly they agreed or disagreed with each statement on a six-point scale ranging from “definitely disagree” (1) to “definitely agree” (6). Extreme response bias was calculated as the proportion of items to which each participant responded with endpoint categories (i.e., the first or sixth interval).

Statistical Analyses

The relationships between the variables of interest were examined by calculating partial correlations, controlling for the age, gender, social desirability bias, and extreme response bias of the participants. Statistical significance was set at 5%. These analysis strategies were preregistered. Uncorrected *p*-values are reported unless otherwise noted, as our sample size calculation was based on the uncorrected significance level. Although not preregistered, we also report the results of significance tests corrected for the false discovery rate (FDR) using the Benjamini and Hochberg (1995) procedure (Groppe et al., 2011). Pearson’s correlation coefficients were calculated in most analyses. When the distribution of a variable was found to be highly skewed, Spearman’s rank correlation coefficients were calculated. Data from two UK participants who selected “other” for their gender were excluded from

the calculation of partial correlations.

Results

Reading Habits

Table 1 presents the descriptive statistics for the ART score and the self-report measures of reading habits (see Supplementary Figure 2 for a visualization of the distributions of these variables), as well as the results of statistical tests of mean differences between the UK and Japanese samples. Note that the distributions of self-reported numbers of books read per year were highly skewed. Overall, UK participants tended to self-report reading more books and engaging in more reading-related activities than Japanese participants. Although the mean ART score was significantly higher for Japanese participants than for UK participants, this difference was difficult to interpret because different versions of the ART were used for the two samples.

Table 2 shows the partial correlations between the ART score and the self-report measures of reading habits. Partial Spearman's rank correlations were computed to account for the highly skewed distributions of self-reports of the number of books read per year (see Table 1 and Supplementary Figure 2). The correlations were generally positive, ranging from $-.059$ to $.504$ in the UK ($Mdn = .278$; 11 out of 13 correlations were significantly positive at both the uncorrected and FDR-corrected levels) and from $.111$ to $.416$ in Japan ($Mdn = .240$; 12 out of 13 correlations were significantly positive at both the uncorrected and FDR-corrected levels). The ART score had the highest correlation with reading literary fiction in the UK and with reading suspense in Japan, which was significantly greater than

the correlation with reading nonfiction in each sample (UK: $t(292) = 2.706$, $p = .007$; Japan: $t(294) = 2.076$, $p = .039$). To account for possible inaccuracies in self-reported numbers of books read per year, especially among avid readers, we also transformed these variables into seven ordinal categories (i.e., 0, 1, ..., 5, and 6, where “6” means “6 books or more”) and examined their correlations with the ART score. The correlations were significant at the FDR-corrected level, except for those for reading romance in the UK and books other than fiction and nonfiction in both countries (see Supplementary Table 3). Although the accuracy of self-reported reading habits may be limited, the results support the use of the ART score as an index of fiction reading habits. A comparison of the correlations between the two samples shows that the associations of the ART score with reading classics and visiting a bookstore were significantly stronger in the UK than in Japan, while its associations with reading romance and suspense were significantly stronger in Japan than in the UK. They were significant at the FDR-corrected levels, except for the difference in visiting a bookstore.

Social Cognition

Table 1 presents the descriptive statistics for the variables related to social cognition, as well as the results of statistical tests of mean differences between the UK and Japanese samples. Compared to Japanese participants, UK participants scored significantly higher on the RMET, the basic emotion recognition task, and the FBTI task, and significantly lower on the Physiognomic Belief Scale and the ASQ. However, these mean differences must be interpreted with caution, as the scalar invariance of these measures between the two countries is unknown (Gregorich, 2006).

Table 3 shows the partial Pearson's correlations between the variables related to social cognition. Eight (at the uncorrected level) or seven (at the FDR-corrected level) of the total 15 correlations were statistically significant for UK participants and seven (at both the uncorrected and FDR-corrected levels) were statistically significant for Japanese participants, six of which were common to both samples. The common correlations were theoretically meaningful, demonstrating associations between different measures of mentalizing skills, between different measures of stereotyping, and between FBTI and facial emotion recognition (Suzuki et al., 2022). At the uncorrected level, Japanese participants showed statistically higher correlations between scores on the RMET and the ASQ, between scores on the FBTI task and Jaeger et al.'s (2020a) questionnaire on physiognomic belief, and between scores on the two questionnaires on physiognomic belief than the UK participants. However, these differences were insignificant at the FDR-corrected level, and the overall pattern of correlations was similar between the UK and Japan, suggesting that the variables related to social cognition represent a similar set of constructs in the two countries.

Relationships Between Fiction Reading Habits and Social Cognition

Table 4 shows the partial Pearson's correlations between the ART score and the variables related to social cognition (see Figure 1 for a visualization of the relationships). In the UK sample, the ART score had significant positive correlations with the scores on the mentalizing tasks (i.e., the RMET and the basic emotion recognition task). The ART score also had significant negative correlations with some measures of stereotyping (i.e., the scores on Jaeger et al.'s (2020a) questionnaire on

physiognomic belief and the ASQ). These correlations were significant at the FDR-corrected level, except for the one with the physiognomic belief questionnaire. In contrast, no correlation reached statistical significance in Japan. The correlations of the ART score with scores on the RMET and the ASQ were statistically significantly different between the UK and Japan at the uncorrected level; the latter difference was significant at the FDR-corrected level.

Reading Habits of Literary and Popular Fiction

Non-preregistered exploratory analyses were conducted to examine whether reading habits of literary and popular fiction have different relationships with social cognition. The UK version of the ART included a similar number of authors of literary and popular fiction ($n \geq 10$ for each; see below for details) according to the criteria used in previous studies (Black et al., 2018; Mar et al., 2006). However, there seems to be an extreme imbalance in the numbers of the two types of fiction authors in the Japanese version of the ART. Of the 52 authors, three won the Akutagawa Prize, a prestigious award for “pure literature” (*junbungaku*) in Japan, while 32 won the Naoki Prize, a prestigious award for “entertainment literature” in Japan (checked by the first author on August 21, 2021 via <https://www.bunshun.co.jp/shinkoukai/award/>). Although preliminary, this suggests only 3 literary fiction authors and as many as 32 popular fiction authors in the Japanese version of the ART. Therefore, the following analyses were limited to data from UK participants.

First, literary and popular fiction authors were extracted from the authors in the UK version of the ART using a list of literary and popular fiction authors (classified based on publication records)

provided by Mar et al. (2006) and Black et al. (2018). This procedure led to the identification of 10 literary and 13 popular fiction authors (Supplementary Table 1). Second, for each participant, the proportion of correctly identified literary fiction authors minus the proportion of incorrectly identified non-authors was calculated and referred to as literary ART scores. Similarly, popular ART scores were calculated based on the proportion of correctly identified popular fiction authors.

Table 5 shows the partial Spearman's rank correlations between the two types of ART scores and the self-reported numbers of books of different types read per year. The literary ART score was positively correlated with the number of literary fiction books (classics and modern/contemporary literary fiction) read annually to a greater extent than the popular ART score. The difference regarding modern/contemporary literary fiction was significant at the FDR-corrected level. The results suggest that the literary ART score is a better index of literary fiction reading habits than the popular ART score.

Table 5 also shows the partial Pearson's correlations between the two types of ART scores and the variables related to social cognition. Similar to the original ART scores (Table 4), both the literary and popular ART scores had positive correlations with the scores on the mentalizing tasks, and a negative correlation with the score on the ASQ. The popular ART score was also negatively associated with the score on Jaeger et al.'s (2020a) questionnaire on physiognomic belief. These correlations were significant at the FDR-corrected level. At the uncorrected level, the magnitude of the negative correlation with acceptance of stereotyping was greater for the literary ART score than for the popular ART score.

Discussion

1 The main finding of the present study was that fiction reading habits were associated with a fa-
2 vorable social cognitive profile—higher mentalizing skills and lower stereotyping—in UK adults, but
3 not in Japanese adults. The discrepancy between the UK and Japan highlights the importance of a
4 cross-cultural investigation into the links between fiction reading and social cognition (Kevane, 2020;
5 Tabullo et al., 2018; Takahashi et al., 2023), which has been studied primarily in North American and
6 Western European countries (Mumper & Gerrig, 2017).

7 The positive associations between fiction reading habits and mentalizing skills beyond RMET per-
8 formance in UK adults replicate and extend the findings of previous studies (answer to Question 1).
9 Mumper and Gerrig's (2017) meta-analysis showed an aggregate correlation of $r = .211$ between fic-
10 tion reading habits and mental inference from the eyes, as measured by the ART and RMET, respec-
11 tively. We not only replicated this relationship with a similar effect size ($r = .230$) but also demonstrated
12 an association between fiction reading habits and basic emotion recognition from facial expressions (r
13 $= .184$). Although the association between ART and RMET scores has been interpreted as evidence of
14 a link between fiction reading habits and social cognition (Fong et al., 2013; Kidd & Castano, 2017; Mar
15 et al., 2006, 2009), it may instead reflect the verbal knowledge developed during fiction reading. That
16 is, a rich vocabulary is necessary to score high on the RMET, as this test uses uncommon words to
17 describe mental states (Dodell-Feder et al., 2020). In contrast, the basic emotion recognition task in
18 the present study used only common words (i.e., basic emotion terms) as response options. Thus, the
19 observed association between the scores on this task and ART provides clearer evidence of a

relationship between fiction reading habits and mentalizing skills that is less likely to be confounded by verbal ability.

Regarding Question 2, the associations between fiction reading habits and stereotyping observed in the UK were complicated. On the one hand, fiction reading habits were negatively associated with acceptance of stereotyping (Carter et al., 2006), a general measure of stereotyping. On the other hand, the correlations between fiction reading habits and facial stereotyping were mixed. Specifically, fiction reading habits were negatively correlated with one measure of physiognomic belief (Jaeger et al., 2020a) but not with its other measure (Suzuki et al., 2019) or with a task-based measure of facial stereotyping (FBTI extremity; Suzuki et al., 2019, 2022). Jaeger et al.'s (2020a) questionnaire differs from the other two measures in that it assesses general beliefs about FBTI without reference to specific traits, which may have led to the different statistical results. In any case, the correlations between fiction reading habits and the three measures of facial stereotyping were all small, just above or below .10 in absolute terms. It has been argued that because fiction tends to contain counter-stereotypical portrayals of people and the world, reading fiction can mitigate a range of stereotype-related beliefs (Buttrick et al., 2023; Castano et al., 2020, 2021; Fong et al., 2015). However, fiction writers also tend to describe the facial appearance of story characters in order to give readers specific impressions about them (Kronshage, 2018; Tytler, 1982). Available evidence suggests that intensive exposure to counter-stereotypical exemplars (e.g., learning that untrustworthy-looking people are four times more likely to be trustworthy than untrustworthy in laboratory tasks) may be necessary to reduce facial

1 stereotyping (Chua & Freeman, 2021; Jaeger et al., 2020b). However, people who enjoy reading fiction
2 are likely to read a variety of books, some of which may or may not contain stereotypical descriptions
3 of characters' faces. Such natural exposure to the mixed descriptions of facial stereotyping during lei-
4 sure reading may not be sufficient to change readers' beliefs. Consequently, the relationships between
5 fiction reading habits and facial stereotyping may necessarily be observed as weak or nonexistent. To
6 address this limitation, it would be useful to conduct an experimental study that compares the effects
7 of exposure to fiction books that affirm and disaffirm facial stereotyping.

8 In contrast to the findings in the UK, we observed a general lack of relationships between fiction
9 reading habits and social cognition in Japanese adults (answer to Question 3), which requires special
10 explanation. The results are generally consistent with those of Takahashi et al. (2023), who showed
11 that the correlations between fiction reading habits and RMET performance among Japanese under-
12 graduates are weak at best, ranging from .003 to .110. Here, we discuss two possible explanations for
13 the negative findings in Japan.

14 One possibility is that methodological issues were responsible for the national differences ob-
15 served in this study. First, the UK and Japanese versions of the ART appear to differ in the composition
16 of author types. Specifically, as detailed in the Results section, the UK version (Gilbert et al., 2018)
17 seems to include the names of both literary and popular fiction authors (Black et al., 2018; Mar et al.,
18 2006), whereas the Japanese version (Osanai et al., 2015) seems to include mostly popular fiction au-
19 thors. It has been argued that literary fiction has a greater potential to cultivate social cognition than

popular fiction because complex and subtle descriptions of people are more common in literary fiction than in popular fiction (Buttrick et al., 2023; Castano, 2024; Kidd & Castano, 2013, 2017). Thus, the UK and Japanese versions of the ART may have measured different types of fiction reading habits, which could account for the different relationships between the ART score and social cognitive variables in these two countries. Second, photos of eyes constituting stimuli of the RMET used in this study were those of Caucasian individuals (Baron-Cohen et al., 2001). The difficulty in inferring the mental states of people of other races (Adams et al., 2009) may explain the lower mean and reliability of the RMET scores for the Japanese participants than for the UK participants in the present study (Table 1), which may in turn weaken the relationships with other variables. These methodological factors may have obscured otherwise existing relationships between fiction reading habits and social cognition in Japan.

However, objections can be raised to the methodological explanations. First, it remains controversial whether literary fiction has a greater power to promote social cognition than popular fiction (Fong et al., 2013, 2015; Stern et al., 2019). Based on post hoc and exploratory analyses, we found that familiarity with literary fiction authors and popular fiction authors had comparable associations with social cognitive variables in UK adults. The results would undermine the explanation that, at least for the current data, scores on the UK version of the ART are positively correlated with social cognitive variables because they reflect habits of reading literary fiction rather than popular fiction. Second, although the means of most of the social cognitive variables, including the RMET score, differed between the UK and Japanese participants, each variable had a broad distribution and did not indicate a

1 floor or ceiling effect. Since an association between two variables is invariant to parallel translation,
2 we suggest that the mean difference in the RMET score between the UK and Japan cannot explain the
3 differences in the correlation results. Instead, from a psychometric perspective, the lower reliability of
4 the RMET score among Japanese participants could reduce its correlations with other measures. How-
5 ever, the score on the basic emotion recognition task was differentially associated with fiction reading
6 habits between Japan and the UK, similar to the RMET scores. The basic emotion recognition task used
7 Caucasian and Japanese faces for the UK and Japanese participants, respectively, and the reliability of
8 its scores was comparable between the two samples (Table 1). Therefore, it would be difficult to ex-
9 plain the national differences in the associations between fiction reading habits and mentalizing skills
10 as merely reflecting differences in measurement reliability. Nevertheless, the differences between lit-
11 erary and popular fiction in their relationships with social cognition need to be further explored by
12 studies using ARTs that are designed to include both types of fiction authors in a balanced manner. To
13 avoid possible confounding influences of other-race effects, it is also important to include stimuli from
14 people of the same culture as the participants in social cognitive tasks (Adams et al., 2009).

15 Another possibility is the effect of culture-specific storytelling traditions, if such traditions exist
16 (Buttrick et al., 2023). Fiebich (2016) highlighted the distinction between mentalistic and behavioral-
17 contextual narrative practices. To understand people's behavior, the former prefers explicit references
18 to mental states such as emotions and beliefs, whereas the latter prefers explicit references to socio-
19 situational contexts such as social roles and norms. Cultural differences in the independence and

interdependence of social relationships (Markus & Kitayama, 2010) appear to vary according to preferred narrative practices (Fiebich, 2016; Lillard, 1999; Slaughter & Perez-Zapata, 2014). According to this view, mentalistic narrative practices are more prevalent in North American and Western European countries such as the UK, where independent types of social relations guided by personal preferences and goals are assumed to be common. In contrast, behavioral-contextual narrative practices are more prevalent in East Asian countries such as Japan, where interdependent types of social relations guided by communal and relational concerns are assumed to be common. For example, when justifying responses to false belief tasks, Western and Japanese children tend to mention the mental states of characters and interpersonal rules, respectively (Bartsch & Wellman, 1989; Naito & Koyama, 2006). Western children's mentalistic narrative practices may be inherited from their caregivers: European American mothers in the United States make more references to mental states during storytelling than do Chinese immigrant mothers (Doan & Wang, 2010). If these cultural differences in narrative practices apply to fiction, Western readers may be more exposed to mentalistic representations than Eastern readers, possibly resulting in stronger associations between fiction reading habits and mentalizing skills in Western countries than in Eastern countries.

A major limitation of the present study is the difficulty in determining whether the observed differences between the UK and Japan are due to methodological and/or cultural factors. A methodological explanation may seem unlikely, given that previous studies using methods different from ours have also found non-robust or null associations between fiction reading and social cognition in non-

Western countries (Kevane, 2020; Tabullo et al., 2018; Takahashi et al., 2023). However, evidence is insufficient to draw firm conclusions. Future cross-cultural studies using versions of the ART that carefully control for author genre composition (Black et al., 2018) and include more diverse social cognition tasks would help address this issue. For example, in addition to mentalizing about isolated individuals, such as in the RMET, measuring mental inferences about groups of people may highlight cultural differences (Masuda et al., 2008). Textual analyses that explore culture-specific storytelling traditions in fictional works would also be useful (Imada, 2012). It is also possible that the nature of our FBTI task, which required all participants to make trait inferences from faces alone, may have obscured individual differences in facial stereotyping. Modifying the trait inference task to include multiple cues (e.g., facial appearance and behavior [Suzuki, 2018] or situational information [Jaeger et al., 2019]) may provide a more fine-grained measure of individual differences in FBTI, which may in turn help us detect their relationship with fiction reading habits. In addition, experimental studies that compare the causal effects of reading on social cognition across cultures are important (Dodell-Feder & Tamir, 2018; Kevane, 2020; Kidd & Castano, 2013; Takahashi et al., 2023). In these studies, multilingual translated versions of the same narrative work can be used to control for the confounding influences of (cultural) differences in fictional content.

In conclusion, the present study shows that among UK adults, mentalizing skills positively correlated with fiction reading habits include not only complex mental state inferences from eyes but also basic emotion recognition from facial expressions. The results clarify that the positive association

between fiction reading habits and mentalizing skills is not limited to a specific mentalizing task that requires rich verbal knowledge (RMET). Furthermore, while fiction reading habits were negatively correlated with acceptance of stereotyping in general, their associations with facial stereotyping were mixed. This finding suggests that the relationships between fiction reading habits and stereotyping may differ depending on the content and target of the stereotype. Finally, no statistically significant relationship was found between fiction reading habits and social cognition among Japanese adults. Future research is required to clarify the causes of the observed differences between the UK and Japan.

References

- Adams, R. B., Jr., Rule, N. O., Franklin, R. G., Jr., Wang, E., Stevenson, M. T., Yoshikawa, S., Nomura, M., Sato, W., Kveraga, K., & Ambady, N. (2010). Cross-cultural reading the mind in the eyes: An fMRI investigation. *Journal of Cognitive Neuroscience*, 22(1), 97–108.
<https://doi.org/10.1162/jocn.2009.21187>
- Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The “Reading the Mind in the Eyes” test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry*, 42(2), 241–251.
<https://doi.org/10.1111/1469-7610.00715>
- Bartsch, K., & Wellman, H. (1989). Young children’s attribution of action to beliefs and desires. *Child Development*, 60(4), 946–964. <https://doi.org/10.2307/1131035>

- 1 Batson, C. D., & Ahmad, N. Y. (2009). Using empathy to improve intergroup attitudes and relations.
2 *Social Issues and Policy Review*, 3(1), 141–177. <https://doi.org/10.1111/j.1751->
3 2409.2009.01013.x
- 4 Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful ap-
5 proach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*,
6 57(1), 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- 7 Black, J. E., Capps, S. C., & Barnes, J. L. (2018). Fiction, genre exposure, and moral reality. *Psychology*
8 *of Aesthetics, Creativity, and the Arts*, 12 (3), 328–340. <https://doi.org/10.1037/aca0000116>
- 9 Buttrick, N., Westgate, E. C., & Oishi, S. (2023). Reading literary fiction is associated with a more com-
10 plex worldview. *Personality and Social Psychology Bulletin*, 49(9), 1408–1420.
11 <https://doi.org/10.1177/01461672221106059>
- 12 Calder, A. J., Young, A. W., Rowland, D., Perrett, D. I., Hodges, J. R., & Etcoff, N. L. (1996). Facial emotion
13 recognition after bilateral amygdala damage: Differentially severe impairment of fear. *Cogni-*
14 *tive Neuropsychology*, 13(5), 699–745. <https://doi.org/10.1080/026432996381890>
- 15 Carter, J. D., Hall, J. A., Carney, D. R., & Rosip, J. C. (2006). Individual differences in the acceptance of
16 stereotyping. *Journal of Research in Personality*, 40(6), 1103–1118.
17 <https://doi.org/10.1016/j.jrp.2005.11.005>
- 18 Castano, E. (2024). Less is more: How the language of fiction fosters emotion recognition. *Emotion*
19 *Review*. Advance online publication. <https://doi.org/10.1177/17540739241232350>

- 1 Castano, E., Martingano, A. J., & Perconti, P. (2020). The effect of exposure to fiction on attributional
2 complexity, egocentric bias and accuracy in social perception. *PLOS ONE*, 15(5), Article
3 e0233378. <https://doi.org/10.1371/journal.pone.0233378>
- 4 Castano, E., Paladino, M. P., Cadwell, O. G., Cuccio, V., & Perconti, P. (2021). Exposure to literary fiction
5 is associated with lower psychological essentialism. *Frontiers in Psychology*, 12, Article 662940.
6 <https://doi.org/10.3389/fpsyg.2021.662940>
- 7 Chua, K.-W., & Freeman, J. B. (2021). Facial stereotype bias is mitigated by training. *Social Psychological
8 and Personality Science*, 12(7), 1335–1344. <https://doi.org/10.1177/1948550620972550>
- 9 De Mulder, H. N. M., Hakemulder, F., Klaassen, F., Junge, C. M. M., Hoijtink, H., & van Berkum, J. J. A.
10 (2022). Figuring out what they feel: Exposure to eudaimonic narrative fiction is related to men-
11 talizing ability. *Psychology of Aesthetics Creativity and the Arts*, 16(2), 2242–2258.
12 <https://doi.org/10.1037/aca0000428>
- 13 Djikic, M., Oatley, K., & Moldoveanu, M. C. (2013). Reading other minds: Effects of literature on empa-
14 thy. *Scientific Study of Literature*, 3(1), 28–47.
15 <https://doi.org/https://doi.org/10.1075/ssol.3.1.06dji>
- 16 Doan, S. N., & Wang, Q. (2010). Maternal discussions of mental states and behaviors: Relations to
17 emotion situation knowledge in European American and immigrant Chinese children. *Child
18 Development*, 81(5), 1490–1503. [https://doi.org/https://doi.org/10.1111/j.1467-
19 8624.2010.01487.x](https://doi.org/https://doi.org/10.1111/j.1467-8624.2010.01487.x)

- 1 Dodell-Feder, D., Ressler, K. J., & Germine, L. T. (2020). Social cognition or social class and culture? On
2 the interpretation of differences in social cognitive performance. *Psychological Medicine*,
3 50(1), 133–145. <https://doi.org/10.1017/s003329171800404x>
- 4 Dodell-Feder, D., & Tamir, D. I. (2018). Fiction reading has a small positive impact on social cognition:
5 A meta-analysis. *Journal of Experimental Psychology: General*, 147(11), 1713–1727.
6 <https://doi.org/10.1037/xge0000395>
- 7 Eekhof, L. S., van Krieken, K., & Willems, R. M. (2022). Reading about minds: The social-cognitive po-
8 tential of narratives. *Psychonomic Bulletin & Review*, 29, 1703–1718.
9 <https://doi.org/10.3758/s13423-022-02079-z>
- 10 Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6(3–4), 169–200.
11 <https://doi.org/10.1080/02699939208411068>
- 12 Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect* [Photographs]. Consulting Psychologists
13 Press.
- 14 Ekman, P., Friesen, W. V., & Hager, J. C. (2002). *Facial action coding system: Investigator's guide*. Re-
15 search Nexus.
- 16 Faraway, J. J. (2014). *Linear models with R* (2nd ed.). CRC Press.
- 17 Fiebich, A. (2016). Narratives, culture, and folk psychology. *Phenomenology and the Cognitive Sciences*,
18 15(1), 135–149. <https://doi.org/10.1007/s11097-014-9378-7>

- 1 Fong, K., Mullin, J. B., & Mar, R. A. (2013). What you read matters: The role of fiction genre in predicting
2 interpersonal sensitivity. *Psychology of Aesthetics, Creativity, and the Arts*, 7(4), 370–376.
3 <https://doi.org/10.1037/a0034084>
- 4 Fong, K., Mullin, J. B., & Mar, R. A. (2015). How exposure to literary genres relates to attitudes toward
5 gender roles and sexual behavior. *Psychology of Aesthetics, Creativity, and the Arts*, 9(3), 274–
6 285. <https://doi.org/10.1037/a0038864>
- 7 Gilbert, R. A., Davis, M. H., Gaskell, M. G., & Rodd, J. M. (2018). Listeners and readers generalize their
8 experience with word meanings across modalities. *Journal of Experimental Psychology: Learn-*
9 *ing, Memory, and Cognition*, 44(10), 1533–1561. <https://doi.org/10.1037/xlm0000532>
- 10 Gregorich, S. E. (2006). Do self-report instruments allow meaningful comparisons across diverse pop-
11 ulation groups? Testing measurement invariance using the confirmatory factor analysis frame-
12 work. *Medical Care*, 44(11), S78–S94. <https://doi.org/10.1097/01.mlr.0000245454.12228.8f>
- 13 Greenleaf, E. A. (1992). Measuring extreme response style. *Public Opinion Quarterly*, 56(3), 328–351.
14 <https://doi.org/10.1086/269326>
- 15 Groppe, D. M., Urbach, T. P., & Kutas, M. (2011). Mass univariate analysis of event-related brain poten-
16 tials/fields I: A critical tutorial review. *Psychophysiology*, 48(12), 1711–1725.
17 <https://doi.org/10.1111/j.1469-8986.2011.01273.x>

- 1 Hart, C. M., Ritchie, T. D., Hepper, E. G., & Gebauer, J. E. (2015). The Balanced Inventory of Desirable
2 Responding Short Form (BIDR-16). *SAGE Open*, 5(4), Article 2158244015621113.
3 <https://doi.org/10.1177/2158244015621113>
- 4 Hirao, K., Miyata, J., Fujiwara, H., Yamada, M., Namiki, C., Shimizu, M., Sawamoto, N., Fukuyama, H.,
5 Hayashi, T., & Murai, T. (2008). Theory of mind and frontal lobe pathology in schizophrenia: A
6 voxel-based morphometry study. *Schizophrenia Research*, 105(1), 165–174.
7 <https://doi.org/10.1016/j.schres.2008.07.021>
- 8 Imada, T. (2010). Cultural narratives of individualism and collectivism: A content analysis of textbook
9 stories in the United States and Japan. *Journal of Cross-Cultural Psychology*, 43(4), 576–591.
10 <https://doi.org/10.1177/0022022110383312>
- 11 Jaeger, B., Evans, A. M., Stel, M., & van Beest, I. (2019). Explaining the persistent influence of facial
12 cues in social decision-making. *Journal of Experimental Psychology: General*, 148(6), 1008–
13 1021. <https://doi.org/10.1037/xge0000591>
- 14 Jaeger, B., Evans, A., Stel, M., & van Beest, I. (2020a). *Lay beliefs in physiognomy explain overreliance*
15 *on facial impression*. PsyArXiv. <https://doi.org/10.31234/osf.io/8dq4x>
- 16 Jaeger, B., Todorov, A. T., Evans, A. M., & van Beest, I. (2020b). Can we reduce facial biases? Persistent
17 effects of facial trustworthiness on sentencing decisions. *Journal of Experimental Social Psy-*
18 *chology*, 90, Article 104004. <https://doi.org/https://doi.org/10.1016/j.jesp.2020.104004>

- 1 Johnson, D. R., Jasper, D. M., Griffin, S., & Huffman, B. L. (2013). Reading narrative fiction reduces Arab-
2 Muslim prejudice and offers a safe haven from intergroup anxiety. *Social Cognition*, 31(5), 578–
3 598. <https://doi.org/10.1521/soco.2013.31.5.578>
- 4 Kevane, M. (2020). Reading fiction and economic preferences of rural youth in Burkina Faso. *Economic*
5 *Development & Cultural Change*, 68(3), 1041–1079. <https://doi.org/10.1086/701704>
- 6 Kidd, D. C., & Castano, E. (2013). Reading literary fiction improves theory of mind. *Science*, 342 (6156),
7 377–380. <https://doi.org/10.1126/science.1239918>
- 8 Kidd, D., & Castano, E. (2017). Different stories: How levels of familiarity with literary and genre fiction
9 relate to mentalizing. *Psychology of Aesthetics, Creativity, and the Arts*, 11(4), 474–486.
10 <https://doi.org/10.1037/aca0000069>
- 11 Koopman, E. M., & Hakemulder, F. (2015). Effects of literature on empathy and self-reflection: A theo-
12 retical-empirical framework. *Journal of Literary Theory*, 9(1), 79–111.
13 <https://doi.org/doi:10.1515/jlt-2015-0005>
- 14 Kronshage, E. (2018). *Vision and character: Physiognomics and the English Realist novel*. Routledge.
- 15 Lehne, M., Engel, P., Rohrmeier, M., Menninghaus, W., Jacobs, A. M., & Koelsch, S. (2015). Reading a
16 suspenseful literary text activates brain areas related to social cognition and predictive infer-
17 ence. *PLOS ONE*, 10(5), Article e0124550. <https://doi.org/10.1371/journal.pone.0124550>

- 1 Lenhart, J., Dangel, J., & Richter, T. (2022). The relationship between lifetime book reading and empa-
2 thy in adolescents: Examining transportability as a moderator. *Psychology of Aesthetics, Crea-*
3 *tivity, and the Arts*, 16(4), 679–693. <https://doi.org/10.1037/aca0000341>
- 4 Lillard, A. (1999). Developing a cultural theory of mind: The CIAO approach. *Current Directions in Psy-*
5 *chological Science*, 8(2), 57–61. <https://doi.org/10.1111/1467-8721.00014>
- 6 Livingston, R. W. (2001). What you see is what you get: Systematic variability in perceptual-based social
7 judgment. *Personality and Social Psychology Bulletin*, 27(9), 1086–1096.
8 <https://doi.org/10.1177/0146167201279002>
- 9 Lundqvist, D., Flykt, A., & Öhman, A. (1998). *Karolinska Directed Emotional Faces (KDEF)* [Photographs].
10 Karolinska Institutet.
- 11 Majid, A. (2023). Establishing psychological universals. *Nature Reviews Psychology*, 2(4), 199–200.
12 <https://doi.org/10.1038/s44159-023-00169-w>
- 13 Mar, R. A. (2018). Evaluating whether stories can promote social cognition: Introducing the Social Pro-
14 cesses and Content Entrained by Narrative (SPaCEN) framework. *Discourse Processes*, 55(5–6),
15 454–479. <https://doi.org/10.1080/0163853X.2018.1448209>
- 16 Mar, R. A., & Oatley, K. (2008). The function of fiction is the abstraction and simulation of social expe-
17 rience. *Perspectives on Psychological Science*, 3(3), 173–192. [https://doi.org/10.1111/j.1745-](https://doi.org/10.1111/j.1745-6924.2008.00073.x)
18 [6924.2008.00073.x](https://doi.org/10.1111/j.1745-6924.2008.00073.x)

- 1 Mar, R. A., Oatley, K., Hirsh, J., dela Paz, J., & Peterson, J. B. (2006). Bookworms versus nerds: Exposure
2 to fiction versus non-fiction, divergent associations with social ability, and the simulation of
3 fictional social worlds. *Journal of Research in Personality*, 40(5), 694–712.
4 <https://doi.org/10.1016/j.jrp.2005.08.002>
- 5 Mar, R. A., Oatley, K., & Peterson, J. B. (2009). Exploring the link between reading fiction and empathy:
6 Ruling out individual differences and examining outcomes. *Communications*, 34(4), 407–428.
7 <https://doi.org/10.1515/COMM.2009.025>
- 8 Markus, H. R., & Kitayama, S. (2010). Cultures and selves: A cycle of mutual constitution. *Perspectives*
9 *on Psychological Science*, 5(4), 420–430. <https://doi.org/10.1177/1745691610375557>
- 10 Martinez, A. M. (2017). Visual perception of facial expressions of emotion. *Current Opinion in Psychol-*
11 *ogy*, 17, 27–33. <https://doi.org/https://doi.org/10.1016/j.copsyc.2017.06.009>
- 12 Masuda, T., Ellsworth, P. C., Mesquita, B., Leu, J., Tanida, S., & Van de Veerdonk, E. (2008). Placing the
13 face in context: Cultural differences in the perception of facial emotion. *Journal of Personality*
14 *and Social Psychology*, 94(3), 365–381. <https://doi.org/10.1037/0022-3514.94.3.365>
- 15 Moore, M., & Gordon, P. C. (2015). Reading ability and print exposure: item response theory analysis
16 of the author recognition test. *Behavior Research Methods*, 47(4), 1095–1109.
17 <https://doi.org/10.3758/s13428-014-0534-3>
- 18 Mumper, M. L., & Gerrig, R. J. (2017). Leisure reading and social cognition: A meta-analysis. *Psychology*
19 *of Aesthetics, Creativity, and the Arts*, 11(1), 109–120. <https://doi.org/10.1037/aca0000089>

- 1 Naito, M., & Koyama, K. (2006). The development of false-belief understanding in Japanese children:
2 Delay and difference? *International Journal of Behavioral Development*, 30(4), 290–304.
3 <https://doi.org/10.1177/0165025406063622>
- 4 Oakley, B. F. M., Brewer, R., Bird, G., & Catmur, C. (2016). Theory of mind is not theory of emotion: A
5 cautionary note on the Reading the Mind in the Eyes Test. *Journal of Abnormal Psychology*,
6 125(6), 818–823. <https://doi.org/10.1037/abn0000182>
- 7 Oatley, K., & Djikic, M. (2018). Psychology of narrative art. *Review of General Psychology*, 22(2), 161–
8 168. <https://doi.org/10.1037/gpr0000113>
- 9 Olivola, C. Y., Funk, F., & Todorov, A. (2014). Social attributions from faces bias human choices. *Trends*
10 *in Cognitive Sciences*, 18(11), 566–570. <https://doi.org/10.1016/j.tics.2014.09.007>
- 11 Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the Na-*
12 *tional Academy of Sciences of the United States of America*, 105(32), 11087–11092.
13 <https://doi.org/10.1073/pnas.0805664105>
- 14 Osanai, H., Komeda, H., Furumi, F., & Kusumi, T. (2015, September 22–24). *Measuring reading amount*
15 *in Japanese undergraduates: Development of Japanese version of the Author Recognition Test*
16 [Poster presentation]. The 79th Annual Convention of the Japanese Psychological Association,
17 Nagoya, Aichi, Japan. https://doi.org/10.4992/pacjpa.79.0_2ev-096
- 18 Panero, M. E., Weisberg, D. S., Black, J., Goldstein, T. R., Barnes, J. L., Brownell, H., & Winner, E. (2016).
19 Does reading a single passage of literary fiction really improve theory of mind? An attempt at

replication. *Journal of Personality and Social Psychology*, 111(5), e46–e54.

<https://doi.org/10.1037/pspa0000064>

Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S.

Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17–59). Ac-

ademic Press. <https://doi.org/10.1016/B978-0-12-590241-0.50006-X>

Peterson, E., & Miller, S. (2012). The Eyes Test as a measure of individual differences: How much of the

variance reflects verbal IQ? *Frontiers in Psychology*, 3, Article 220.

<https://doi.org/10.3389/fpsyg.2012.00220>

Quinlan, J. A., Padgett, J. K., Khajehnassiri, A., & Mar, R. A. (2023). Does a brief exposure to literary

fiction improve social ability? Assessing the evidential value of published studies with a p-

curve. *Journal of Experimental Psychology: General*, 152(3), 723–732.

<https://doi.org/10.1037/xge0001302>

Samur, D., Tops, M., & Koole, S. L. (2018). Does a single session of reading literary fiction prime en-

hanced mentalising performance? Four replication experiments of Kidd and Castano (2013).

Cognition and Emotion, 32(1), 130–144. <https://doi.org/10.1080/02699931.2017.1279591>

Schwering, S. C., Ghaffari-Nikou, N. M., Zhao, F., Niedenthal, P. M., & MacDonald, M. C. (2021). Explor-

ing the relationship between fiction reading and emotion recognition. *Affective Science*, 2(2),

178–186. <https://doi.org/10.1007/s42761-021-00034-0>

- 1 Slaughter, V., & Perez-Zapata, D. (2014). Cultural variations in the development of mind reading. *Child*
2 *Development Perspectives*, 8(4), 237–241. <https://doi.org/10.1111/cdep.12091>
- 3 Sprengelmeyer, R., Young, A. W., Calder, A. J., Karnat, A., Lange, H., Hömberg, V., Perrett, D. I., & Row-
4 land, D. (1996). Loss of disgust: Perception of faces and emotions in Huntington's disease.
5 *Brain*, 119(5), 1647–1665. <https://doi.org/10.1093/brain/119.5.1647>
- 6 Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research*
7 *Quarterly*, 24(4), 402–433. <https://doi.org/10.2307/747605>
- 8 Stansfield, J., & Bunce, L. (2014). The relationship between empathy and reading fiction: Separate roles
9 for cognitive and affective components. *Journal of European Psychology Students*, 5(3), 9–18.
10 <https://doi.org/10.5334/jeps.ca>
- 11 Stern, S. C., Robbins, B., Black, J. E., & Barnes, J. L. (2019). What you read and what you believe: Genre
12 exposure and beliefs about relationships. *Psychology of Aesthetics, Creativity, and the Arts*,
13 13(4), 450–461. <https://doi.org/10.1037/aca0000189>
- 14 Sutherland, C. A. M., Oldmeadow, J. A., & Young, A. W. (2016). Integrating social and facial models of
15 person perception: Converging and diverging dimensions. *Cognition*, 157, 257–267.
16 <https://doi.org/10.1016/j.cognition.2016.09.006>
- 17 Suzuki, A. (2018). Persistent reliance on facial appearance among older adults when judging someone's
18 trustworthiness. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*,
19 73(4), 573–583. <https://doi.org/10.1093/geronb/gbw034>

- 1 Suzuki, A., Hoshino, T., & Shigemasu, K. (2006a). Measuring individual differences in sensitivities to
2 basic emotions in faces. *Cognition*, 99(3), 327–353. <https://doi.org/10.1016/j.cognition.2005.04.003>
3
- 4 Suzuki, A., Hoshino, T., & Shigemasu, K. (2010). Happiness is unique: A latent structure of emotion
5 recognition traits revealed by statistical model comparison. *Personality and Individual Differences*, 48(2), 196–201. <https://doi.org/10.1016/j.paid.2009.10.006>
6
- 7 Suzuki, A., Hoshino, T., Shigemasu, K., & Kawamura, M. (2006b). Disgust-specific impairment of facial
8 expression recognition in Parkinson's disease. *Brain*, 129(3), 707–717.
9 <https://doi.org/10.1093/brain/awl011>
- 10 Suzuki, A., Liu, C., & Osanai, H. (2024, July 26). *UK-Japan Binational Study of Correlates of Face-Based*
11 *Trait Inference and Fiction Reading Habits*. Retrieved from osf.io/ntq2m
- 12 Suzuki, A., Tsukamoto, S., & Takahashi, Y. (2019). Faces tell everything in a just and biologically deter-
13 mined world: Lay theories behind face reading. *Social Psychological and Personality Science*,
14 10(1), 62–72. <https://doi.org/10.1177/1948550617734616>
- 15 Suzuki, A., Tsukamoto, S., & Takahashi, Y. (2022). Generalized tendency to make extreme trait judge-
16 ments from faces. *Royal Society Open Science*, 9(11), Article 220172.
17 <https://doi.org/10.1098/rsos.220172>

- 1 Tabullo, A. J., Navas Jiménez, V. A., & García, C. S. (2018). Associations between fiction reading, trait
2 empathy and theory of mind ability. *International Journal of Psychology & Psychological Ther-*
3 *apy*, 18(3), 357–370.
- 4 Takahashi, Y., Himichi, T., Masuchi, A., Nakanishi, D., & Ohtsubo, Y. (2023). Is reading fiction associated
5 with a higher mind-reading ability? Two conceptual replication studies in Japan. *PLOS ONE*,
6 18(6), Article e0287542. <https://doi.org/10.1371/journal.pone.0287542>
- 7 Tamir, D. I., Bricker, A. B., Dodell-Feder, D., & Mitchell, J. P. (2016). Reading fiction and reading minds:
8 The role of simulation in the default network. *Social Cognitive and Affective Neuroscience*,
9 11(2), 215–224. <https://doi.org/10.1093/scan/nsv114>
- 10 Tani, I. (2008). Development of Japanese Version of Balanced Inventory of Desirable Responding (BIDR-
11 J). *Japanese Journal of Personality*, 17(1), 18–28. <https://doi.org/10.2132/personality.17.18>
- 12 Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from faces: De-
13 terminants, consequences, accuracy, and functional significance. *Annual Review of Psychology*,
14 66(1), 519–545. <https://doi.org/10.1146/annurev-psych-113011-143831>
- 15 Tytler, G. (1982). *Physiognomy in the European novel: Faces and fortunes*. Princeton University Press.
- 16 Vezzali, L., Stathi, S., Giovannini, D., Capozza, D., & Trifiletti, E. (2015). *The greatest magic of Harry*
17 *Potter: Reducing prejudice*. *Journal of Applied Social Psychology*, 45(2), 105–121.
18 <https://doi.org/10.1111/jasp.12279>

- 1 Young, A. W., Rowland, D., Calder, A. J., Etcoff, N. L., Seth, A., & Perrett, D. I. (1997). Facial expression
- 2 megamix: Tests of dimensional and category accounts of emotion recognition. *Cognition*, 63(3),
- 3 271–313. [https://doi.org/10.1016/S0010-0277\(97\)00003-6](https://doi.org/10.1016/S0010-0277(97)00003-6)

1 **Table 1**

2 Descriptive and Inferential Statistics of the Main Variables

Variable	UK					Japan					Difference			
	<i>M</i>	<i>SD</i>	Skewness	Kurtosis	α	<i>M</i>	<i>SD</i>	Skewness	Kurtosis	α	<i>t</i> / <i>Z</i> ^c	<i>p</i>	<i>p</i> _{FDR}	<i>r</i> ^d
Reading habits														
ART (min: −1, max: 1)	0.281	0.185	0.925	0.437	-	0.341	0.220	0.466	-0.636	-	<i>t</i> = -3.654	< .001	-	-.148
Number of books read per year (min: 0, max: 100)														
Classics	1.097	2.019	3.261	12.913	-	0.256	1.151	6.751	50.748	-	<i>Z</i> = 9.630	< .001	< .001	.359
Modern/contemporary literary fiction	3.157	6.507	4.587	26.627	-	1.787	4.654	5.768	44.934	-	<i>Z</i> = 4.789	< .001	< .001	.210
Popular fiction: Romance	1.487	3.858	5.388	35.539	-	1.136	3.850	8.262	89.673	-	<i>Z</i> = 2.709	.007	.010	.109
Popular fiction: Fantasy	1.350	3.137	4.127	20.647	-	1.233	4.052	7.904	78.916	-	<i>Z</i> = 1.359	.174	.174	.055
Popular fiction: Science fiction	1.113	2.827	4.619	25.387	-	0.884	4.814	14.848	239.511	-	<i>Z</i> = 2.105	.035	.045	.080
Popular fiction: Suspense/thriller/horror/mystery	2.623	5.269	4.987	34.711	-	2.595	7.301	6.273	50.963	-	<i>Z</i> = 3.210	.001	.002	.142
Popular fiction: Other genres	1.340	2.835	3.977	19.794	-	1.545	6.519	12.111	172.798	-	<i>Z</i> = 1.603	.109	.123	.066
Nonfiction	3.273	4.433	2.159	5.265	-	2.954	9.938	7.323	62.685	-	<i>Z</i> = 6.143	< .001	< .001	.278
Others	0.553	1.459	3.907	18.218	-	2.226	7.004	9.654	124.923	-	<i>Z</i> = -4.190	< .001	< .001	-.157
Reading-related activities (min: 0, max: 4)														
Visiting a bookstore	0.990	0.652	0.082	-0.427	-	1.375	0.842	0.243	-0.514	-	<i>Z</i> = -5.690	< .001	< .001	-.245
Visiting a library	0.840	0.866	0.897	0.466	-	0.787	0.917	0.973	0.122	-	<i>Z</i> = 1.165	.244	.244	.051
Browsing an online bookseller	1.543	0.962	0.339	-0.282	-	1.106	1.126	0.810	-0.312	-	<i>Z</i> = 5.832	< .001	< .001	.264
Checking out book reviews	1.200	0.964	0.576	-0.120	-	0.837	0.918	0.995	0.305	-	<i>Z</i> = 5.019	< .001	< .001	.224
Social cognition														

RMET (min: 0, max: 1)	0.744	0.116	-0.769	0.865	.642	0.599	0.104	-0.875	1.954	.512	$t = 16.189$	< .001	< .001	.552
Basic emotion recognition (min: -5, max: 5)	1.721	0.468	-0.818	2.105	.617	1.469	0.390	-0.072	0.059	.619	$t = 7.197$	< .001	< .001	.282
FBTI (min: 1, max: 6)	4.640	0.432	-0.073	-0.136	.796	4.418	0.440	0.233	-0.307	.823	$t = 6.263$	< .001	< .001	.248
Physiognomic belief 1 ^a (min: 1, max: 4)	2.435	0.545	-0.304	0.223	.890	2.611	0.502	-0.192	0.606	.904	$t = -4.124$	< .001	< .001	-.166
Physiognomic belief 2 ^b (min: 1, max: 7)	3.834	1.164	-0.096	-0.630	.816	3.647	1.187	-0.100	-0.709	.791	$t = 1.958$	.051	.051	.080
Acceptance of stereotyping (min: 1, max: 5)	3.244	0.785	-0.374	-0.074	.850	3.746	0.424	0.246	0.726	.601	$t = -9.757$	< .001	< .001	-.371
Response bias														
Social desirability (min:0, max: 1)	0.303	0.235	0.806	0.019	.730	0.274	0.209	0.766	0.226	.760	$t = 1.606$	.109	.109	.066
Extreme response (min: 0, max: 1)	0.230	0.174	0.795	0.509	.682	0.179	0.177	1.112	0.599	.764	$t = 3.580$	< .001	.001	.145

1 **Note.** The numbers in parentheses in the leftmost “Variable” column represent possible minimum (min) and maximum (max) values for each variable. FDR

2 corrections were performed per group of variables (number of books read per year, reading-related activities, social cognition, and response bias). ART =

3 Author Recognition Test; RMET = “Reading the Mind in the Eyes” Test; FBTI = Face-based trait inference; α = Cronbach’s α ; p_{FDR} = FDR-corrected p -value. ^a

4 Measured using the Physiognomic Belief Scale (Suzuki et al., 2019). ^b Measured using Jaeger et al.’s (2020a) questionnaire. ^c Test statistics (t : Welch’s t -test, Z :

5 Wilcoxon-Mann-Whitney U test) for the differences between the UK and Japan samples. ^d Point biserial or rank biserial correlations representing effect sizes

6 for differences between the UK and Japan samples. Differences in scores related to the ART, basic emotion recognition, FBTI, and social desirability should

7 be interpreted with caution because the English and Japanese versions of the tests consisted of different items.

1 **Table 2**2 *Partial Spearman's Rank Correlations (With 95% CIs) Between the Author Recognition Test Score and the Self-Report Measures of Reading Habits*

Variable	UK	Japan	Difference		
			<i>Z</i>	<i>p</i>	<i>p_{FDR}</i>
Number of books read per year					
Classics	.403 [.302, .494]	.218 [.107, .324]	2.480	.013	.042
Modern/contemporary literary fiction	.504 [.414, .585]	.378 [.276, .472]	1.900	.057	.128
Popular fiction: Romance	-.059 [-.172, .056]	.220 [.109, .326]	3.423	.001	.009
Popular fiction: Fantasy	.182 [.069, .290]	.165 [.052, .274]	0.212	.832	.936
Popular fiction: Science fiction	.335 [.229, .432]	.212 [.100, .318]	1.613	.107	.193
Popular fiction: Suspense/thriller/horror/mystery	.235 [.124, .340]	.416 [.317, .505]	2.452	.014	.042
Popular fiction: Other genres	.278 [.169, .380]	.282 [.174, .384]	0.053	.958	.958
Nonfiction	.352 [.248, .448]	.280 [.171, .382]	0.972	.331	.497
Others	.082 [-.033, .194]	.111 [-.003, .222]	0.357	.721	.927
Reading-related activities					
Visiting a bookstore	.395 [.294, .487]	.228 [.117, .333]	2.253	.024	.096
Visiting a library	.224 [.113, .330]	.309 [.202, .408]	1.107	.268	.326
Browsing an online bookseller	.379 [.277, .473]	.240 [.130, .345]	1.859	.063	.126
Checking out book reviews	.264 [.154, .367]	.337 [.233, .435]	0.982	.326	.326

3 *Note.* Statistically significant correlations and differences are shown in bold at the uncorrected 5% level and are underlined at the FDR-corrected 5% level.

4 FDR corrections were performed per country and per group of variables (number of books read per year and reading-related activities). *p_{FDR}* = FDR-corrected

5 *p*-value.

1 **Table 3**2 *Partial Pearson's Correlations (With 95% CIs) Between the Variables Related to Social Cognition*

Variable pair		UK	Japan	Difference		
				<i>Z</i>	<i>p</i>	<i>p</i> _{FDR}
RMET	Basic emotion recognition	.373 [.270, .467]	.289 [.181, .390]	1.140	.254	.635
	FBTI	.115 [.001, .226]	.071 [-.043, .183]	0.536	.592	.683
	Physiognomic belief 1 ^a	.080 [-.035, .192]	-.032 [-.145, .082]	1.353	.176	.528
	Physiognomic belief 2 ^b	-.055 [-.168, .059]	-.001 [-.115, .112]	0.649	.516	.673
	Acceptance of stereotyping	-.094 [-.206, .021]	.112 [-.002, .223]	2.496	.013	.172
Basic emotion recognition	FBTI	.141 [.027, .251]	.252 [.142, .356]	1.399	.162	.528
	Physiognomic belief 1 ^a	.026 [-.088, .140]	.027 [-.087, .140]	0.008	.993	.993
	Physiognomic belief 2 ^b	-.015 [-.129, .099]	.010 [-.104, .124]	0.305	.760	.814
	Acceptance of stereotyping	-.033 [-.147, .081]	.025 [-.089, .138]	0.701	.484	.673
FBTI	Physiognomic belief 1 ^a	.218 [.106, .324]	.166 [.053, .275]	0.648	.517	.673
	Physiognomic belief 2 ^b	-.019 [-.133, .095]	.167 [.054, .275]	2.272	.023	.172
	Acceptance of stereotyping	.155 [.042, .265]	.102 [-.012, .213]	0.658	.510	.673
Physiognomic belief 1 ^a	Physiognomic belief 2 ^b	.490 [.398, .572]	.608 [.531, .675]	2.052	.040	.200
	Acceptance of stereotyping	.254 [.144, .358]	.206 [.094, .312]	0.616	.538	.673
Physiognomic belief 2 ^b	Acceptance of stereotyping	.263 [.154, .367]	.196 [.084, .303]	0.864	.387	.673

3 *Note.* Statistically significant correlations and differences are shown in bold at the uncorrected 5% level and are underlined at the FDR-corrected 5% level.4 FDR corrections were performed per country. *p*_{FDR} = FDR-corrected *p*-value; RMET = “Reading the Mind in the Eyes” Test; FBTI = Face-based trait inference. ^a5 Measured using the Physiognomic Belief Scale (Suzuki et al., 2019). ^b Measured using Jaeger et al.’s (2020a) questionnaire.

1 **Table 4**2 *Partial Pearson's Correlations (With 95% CIs) Between the Author Recognition Test Score and the Variables Related to Social Cognition*

Variable	UK	Japan	Difference		
			<i>Z</i>	<i>p</i>	<i>p</i> _{FDR}
RMET	<u>.230</u> [.119, .335]	.067 [-.048, .179]	2.028	.043	.129
Basic emotion recognition	<u>.184</u> [.072, .292]	.054 [-.060, .167]	1.599	.110	.220
FBTI	-.086 [-.198, .029]	.015 [-.099, .129]	1.222	.222	.317
Physiognomic belief 1 ^a	-.054 [-.167, .061]	-.035 [-.149, .079]	0.227	.820	.820
Physiognomic belief 2 ^b	-.122 [-.233, -.008]	-.030 [-.144, .084]	1.117	.264	.317
Acceptance of stereotyping	<u>-.218</u> [-.324, -.106]	.071 [-.043, .184]	<u>3.544</u>	< .001	< .001

3 *Note.* Statistically significant correlations and differences are shown in bold at the uncorrected 5% level and are underlined at the FDR-corrected 5% level.4 FDR corrections were performed per country. *p*_{FDR} = FDR-corrected *p*-value; RMET = “Reading the Mind in the Eyes” Test; FBTI = Face-based trait inference. ^a5 Measured using the Physiognomic Belief Scale (Suzuki et al., 2019). ^b Measured using Jaeger et al.’s (2020a) questionnaire.

Table 5

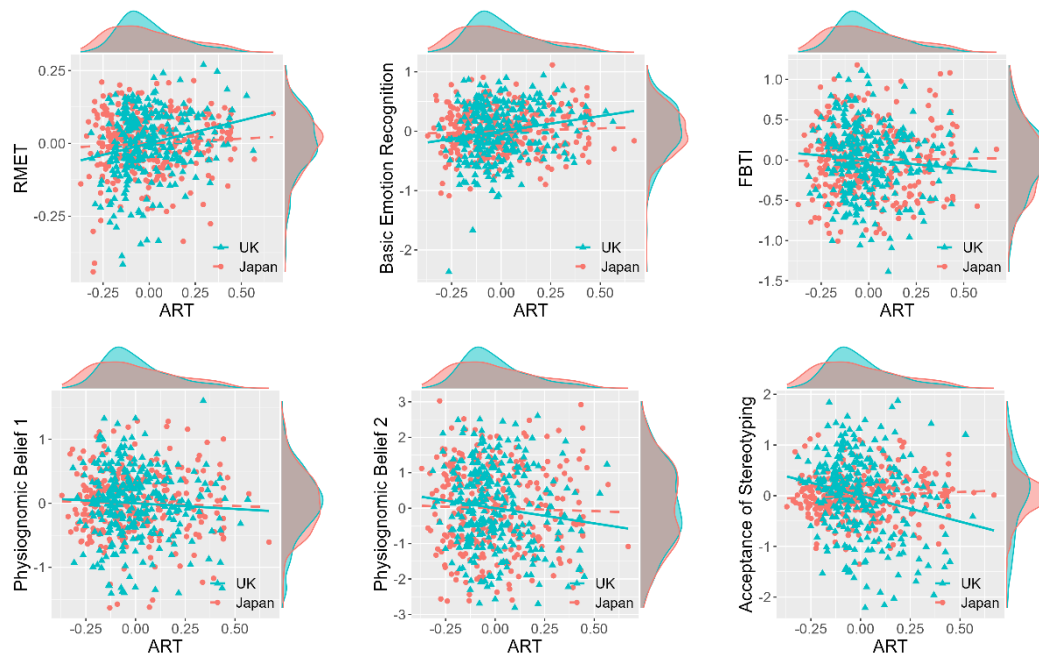
Partial Spearman's Rank Correlations and Pearson's Correlations (With 95% CIs) Between the Two Types of ART Scores and the Self-Reported Numbers of Books Read per Year and the Variables Related to Social Cognition

Variables	Literary ART score	Popular ART score	Difference		
			<i>t</i>	<i>P</i>	<i>p</i> _{FDR}
Number of books read per year					
Classics	<u>.382</u> [.280, .475]	<u>.274</u> [.165, .377]	2.472	.014	.063
Modern/contemporary literary fiction	<u>.479</u> [.386, .563]	<u>.334</u> [.229, .432]	3.512	.001	.009
Popular fiction: Romance	-.109 [-.221, .005]	-.033 [-.147, .081]	-1.637	.103	.309
Popular fiction: Fantasy	<u>.152</u> [.039, .262]	<u>.118</u> [.004, .229]	0.744	.457	.617
Popular fiction: Science fiction	<u>.272</u> [.163, .375]	<u>.304</u> [.196, .404]	-0.706	.480	.617
Popular fiction: Suspense/thriller/horror/mystery	<u>.164</u> [.051, .273]	<u>.224</u> [.112, .330]	-1.293	.197	.416
Popular fiction: Other genres	<u>.234</u> [.123, .339]	<u>.231</u> [.120, .336]	0.056	.956	.956
Nonfiction	<u>.340</u> [.235, .437]	<u>.287</u> [.178, .388]	1.201	.231	.416
Others	.045 [-.069, .158]	.027 [-.088, .141]	0.388	.698	.785
Social cognition					
RMET	<u>.226</u> [.115, .332]	<u>.213</u> [.102, .320]	0.269	.788	.968
Basic emotion recognition	<u>.173</u> [.060, .282]	<u>.176</u> [.063, .284]	-0.052	.959	.968
FBTI	-.076 [-.188, .039]	-.077 [-.190, .037]	0.040	.968	.968
Physiognomic belief 1 ^a	-.027 [-.141, .088]	-.066 [-.179, .048]	0.825	.410	.820
Physiognomic belief 2 ^b	-.086 [-.198, .029]	<u>-.126</u> [-.237, -.012]	0.860	.391	.820
Acceptance of stereotyping	<u>-.233</u> [-.338, -.122]	<u>-.133</u> [-.244, -.019]	-2.143	.033	.198

Note. Partial Spearman's rank correlations were computed for the self-reported numbers of books read per year. Partial Pearson's correlations were computed for the variables related to social cognition. Statistically significant correlations and differences are shown in bold at the uncorrected 5% level and are underlined at the FDR-corrected 5% level. FDR corrections were performed per score and per group of variables (number of books read per year and social cognition). *p*_{FDR} = FDR-corrected *p*-value; RMET = "Reading the Mind in the Eyes" Test; FBTI = Face-based trait inference. ^a Measured using the Physiognomic Belief Scale (Suzuki et al., 2019). ^b Measured using Jaeger et al.'s (2020a) questionnaire.

Figure 1

Partial Regression Plots Between the Author Recognition Test (ART) Score and the Variables Related to Social Cognition With Marginal Density Plots



Note. The abscissa represents the residuals from the regression of the ART score on the covariates (i.e., participant age, gender, social desirability bias, and extreme response bias). The ordinate represents the residuals from the regression of a social cognition measure on the same covariates. The plots thus illustrate the marginal relationships between the ART score and the variables related to social cognition after controlling for the covariates (Faraway, 2014). The solid and dashed lines in each plot represent the linear regression results for the UK and Japanese samples, respectively. RMET = “Reading the Mind in the Eyes” Test; FBTI = Face-based trait inference. Physiognomic Belief 1 = Score on the Physiognomic Belief Scale (Suzuki et al., 2019). Physiognomic Belief 2 = Score on Jaeger et al.’s (2020a) questionnaire.