

# Recognition as a Measure of Television Exposure: Multiple Measures and Their Relationship to Theory of Mind

Jessica E. Black and Jennifer L. Barnes  
University of Oklahoma

The purpose of this research was to develop a test, similar to the Author Recognition Test (Stanovich & West, 1989), to assess exposure to popular and award-winning fictional TV shows airing, across platforms, in the past decade. Three versions of the TV Recognition Test were developed for this purpose. The first, featured in Studies 1 and 2, asked participants to choose which of 66 names (including 30 foils) they could identify as belonging to TV show characters. In Study 3, two alternative methods of measuring fictional TV exposure were assessed: one presented participants with character names accompanied by pictures; the second presented participants with a list of show titles. Recognition of TV show titles, but not characters, was related to performance on a theory of mind task. Further analyses revealed that this effect was stronger for award-winning TV shows, and limited to men, for whom the relation between the theory of mind task and show titles was only significant for award-winning shows. These findings are discussed as a guidepost for future research.

## Public Policy Relevance Statement

Accurate assessment of cumulative TV exposure is necessary to test associations between TV-viewing behavior and a variety of attitudes (e.g., toward environmental policy, immigration, gender roles), traits (e.g., personality, life satisfaction), and behaviors (e.g., voting, volunteering). This article presents a method of measurement, with three specific examples, that can be adapted to different research designs and methodological needs.

**Keywords:** Media exposure, Television Recognition Test, fiction, theory of mind, personality

**Supplemental materials:** <http://dx.doi.org/10.1037/ppm0000254.supp>

By exposing readers and viewers to complex social situations and allowing them privileged access to characters' thoughts and emotions, fictional stories may serve as a form of social simulation (Mar, 2018; Mar & Oatley, 2008; Oatley, 2016). Similarly, it has been argued that fiction allows the reader to practice making mental state attributions (Zunshine, 2006) and to see the world from a wide variety of perspectives (Hakemulder, 2000). In line with these theories, multiple studies (Black & Barnes, 2015a; Kidd & Castano, 2013; Mar, Oatley, Hirsh, de la Paz, & Peterson, 2006; Mar, Oatley, & Peterson, 2009; Panero et al., 2016) have found a correlation between cumulative written fiction exposure, as measured by author recognition checklists, and performance on a theory of mind, or emotion-reading, task, the Reading the Mind in

the Eyes Test (RMET; Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001). Other studies have found that one-time exposure to literary fiction facilitates performance on the RMET (Black & Barnes, 2015a; Kidd & Castano, 2013). However, nonreplications of this work suggest that any effects of one-time fiction exposure on social cognition may be fragile and may depend on who is reading and how they are reading, as well as the specific text being read (Panero et al., 2016; Samur, Tops, & Koole, 2018).

Koopman and Hakemulder (2015) emphasized that the power of a story to affect the way that we think may rely on the degree to which that narrative provides gaps for readers to fill. Such gaps provoke a mental stillness that encourages readers to reflect upon themselves and social interactions in general. A variety of scholars have suggested that literary fiction, which is particularly likely to feature ambiguities, characters who can be difficult to understand, and striking and unfamiliar textual features, may be more adept at inviting readers into a conversation with the author and challenging readers' ability to understand social others and social situations (Bortolussi & Dixon, 2003; Koopman & Hakemulder, 2015; van Lissa, Caracciolo, van Duuren, & van Leuven, 2016). Importantly, some features of literary fiction, such as the presence of complex characters, ambiguities and shade of gray, and narrative gaps that the audience must fill, may also be present in award-winning fiction in other media, such as film.

This article was published Online First July 18, 2019.

✉ Jessica E. Black and Jennifer L. Barnes, Department of Psychology, University of Oklahoma.

Data and supplemental materials (measures) are available at <https://bit.ly/2GxJCSx>.

Correspondence concerning this article should be addressed to Jessica E. Black, Department of Psychology, University of Oklahoma, Dale Hall Tower, Room 705, 455 West Lindsey, Norman, OK 73071. E-mail: [jessica.black@ou.edu](mailto:jessica.black@ou.edu)

Although a great deal of research has investigated the relationship between film and/or TV exposure and behavior (Christakis, Zimmerman, DiGiuseppe, & McCarty, 2004; Huesmann, Moise-Titus, Podolski, & Eron, 2003) and executive function (Hummer, Kronenberger, Wang, Anderson, & Mathews, 2014; Nathanson, Aladé, Sharp, Rasmussen, & Christy, 2014), it is only recently that the body of research on fiction and social cognition has been extended to these media. Black and Barnes (2015b) found that individuals randomly assigned to watch an award-winning TV drama, such as *Mad Men* or *West Wing*, performed higher on the RMET than participants randomly assigned to watch a TV documentary. In contrast, Mar, Tackett, and Moore (2010) found a correlation between theory of mind development and preschoolers' exposure to children's storybooks and films, but not to children's TV shows. The discrepancy between the two results may in part be due to measurement: Black and Barnes used an experimental paradigm, whereas Mar and colleagues used checklists to measure exposure to books, film, and TV rather than an experimental paradigm. Most measures of lifetime TV exposure are self-report, and most are simply estimates of time spent viewing. For example, Christakis and colleagues (2004) had mothers estimate the number of hours their children spent watching TV on weekdays and on weekends; Hummer and colleagues (2014) had participants estimate their TV use over the past year as well as fill out daily diaries. Because much of the research focuses on the effects of TV in general or to news or politics specifically (Dilliplane, Goldman, & Mutz, 2013), most measurement of exposure tends to be of hours spent viewing (per a given unit of time) or diaries; fewer studies attempt to pinpoint a single aspect of TV exposure (but see Huesmann et al., 2003, and Potter & Chang, 1990). Huesmann and colleagues (2003) reported a study in which children identified their favorite TV shows from lists of shows that were either violent or not, as well as how often they watched them: A composite score predicted aggressive behavior in young adulthood. Potter and Chang (1990) simply asked participants how much time they had spent viewing different categories of programming (e.g., news, sports, prime time soap operas).

Instrumentalizing exposure to targeted programming facilitates testing specific hypotheses when variables that simply reflect viewing time would not (Webster & Wakshlag, 2013). What is more, people tend to under- or overreport TV-viewing behavior (Wonneberger & Irazoqi, 2016). Here, we sought to assess exposure to fictional TV specifically with a method parallel to one that has been successfully used to measure print fiction exposure: checklists of author names interspersed with foils, or names that are not authors, commonly referred to as Author Recognition Tests (ART) since Stanovich and West (1989) reported their use in studying effects of literacy. There are various such tests that assess lifetime exposure to print media (Acheson, Wells, & MacDonald, 2008; Black, Capps, & Barnes, 2017; Cipielewski & Stanovich, 1992; Fong, Mullin, & Mar, 2013; Mar et al., 2006; Stanovich & West, 1989), but there are no current measures that do the same for exposure to visual media. Here, we present three separate checklists, the TV Character Recognition Test (TCRT), the TV Show Recognition Test (TSRT), and the TV Character Image Recognition Test, all modeled on the original ART developed by Stanovich and West (1989).

The original ART (Stanovich & West, 1989) was created for use in research on literacy skills. Participants are presented with a list

of author names and foils (names that are not authors), and asked to identify any authors that they recognize on the list. Following signal detection logic (see Stanislaw & Todorov, 1999, for an overview), rate of hits (correctly identified authors) is adjusted with rate of misses (incorrectly selected foils) to estimate discrimination of signal (hits) from noise (misses, or selection foils). The amount of foils selected can be used to adjust for guessing, either by subtracting misses from hits for a total score (Acheson et al., 2008; Black & Barnes, 2015a, 2015b), by discarding participants with a high rate of guessing (Djikic, Oatley, & Moldoveanu, 2013a; Kidd & Castano, 2013) or by including number of foils selected as a covariate (Fong et al., 2013). The advantages of using a checklist to assess print exposure include avoiding socially desirable responding and the ability to tailor the instrument to the study design (Stanovich, Cunningham, & West, 1998). For example, Stanovich and colleagues wanted to assess *free* reading, or the books children read for pleasure; they therefore only included authors that were not likely to be included in school curricula. Similarly, Black and colleagues (2017) used checklists to measure exposure to specific fiction genres, such as romance and science fiction, to tests the unique relations between distinct genres and real-world judgment. Different checklists can be used to measure exposure to different media: Stanovich and West (1989) originally presented a Magazine Recognition Test that consisted of titles of real magazines and foils; Mar et al. (2010) used the method to assess exposure to books, film, and TV in preschool children. Scores on the ART have been validated with reference to diaries of reading time as well as observed reading (Stanovich et al., 1998).

The original ART has been adapted many times (Acheson et al., 2008; Mar et al., 2006) and has been put to frequent use in research on the relationship between reading fiction, emotions, and theory of mind (Black & Barnes, 2015a; Djikic, Oatley, & Moldoveanu, 2013b; Kidd & Castano, 2013). Scores on the ART have been shown to have a more reliable association with performance on a theory of mind task (the RMET) than one-time exposure to different types of fiction (Black & Barnes, 2015a; Kidd & Castano, 2013; Panero et al., 2016). Even when short-term exposure to literary fiction has been shown to have an effect on theory of mind skills (Black & Barnes, 2015a; Kidd & Castano, 2013), the effect of lifetime fiction exposure, as assessed by the ART, was nearly three times as large.

Recent research has shown that exposure to visual media, as well as literary fiction, may affect theory of mind. As mentioned earlier, Black and Barnes (2015b) found that watching award-winning TV dramas had a reliable effect on theory of mind ability as assessed by the RMET. However, although this effect was stable across two experiments, the researchers did not control for exposure to TV or film. If, as is the case with written fiction, cumulative exposure to TV shows facilitates performance on the RMET, then, ideally, future research into the effects of TV viewing on theory of mind and other cognitive or emotional aspects of human behavior will control for previous exposure to TV media.

The purpose of this research was to develop a measure of exposure to popular TV series that would capture any potential relation between long-term exposure to televised fiction and theory of mind as measured by the RMET, and to examine its relationship to print exposure, personality, and emotion-reading ability. In Studies 1 and 2, we tested a simple list of characters from popular TV shows (TCRT); our choice of character names reflected the

goal of measuring selective exposure and the corresponding attention paid to the content in question (see Webster & Wakshlag, 2013). Viewing time can include involuntary exposure to programming and advertisements for programming not necessarily related to the targeted content; laboratory experiments such as those reported by Black and Barnes (2015b) are examples of forced choice. We assumed that character names would reflect preferences for shows that were freely chosen. In Study 3 we augmented the original TCRT with images of the characters as well as additional characters from streaming shows. Study 3 also included a checklist of TV shows (TSRT) as a broader measure of fictional TV exposure. Like the ARTs used to assess print exposure, the TSRT assumes that people who recognize shows may not have actually seen the show in question, but will have been exposed to its name when viewing similar programming or when selecting it on streaming services such as Netflix and Amazon Prime, which group shows in similar categories and make suggestions based on past viewing behavior.

An important goal of this research was to capture any relation between selective exposure to fictional TV and RMET scores. Black and Barnes (2015b) reported that participants scored higher on the RMET in the immediate aftermath of exposure to award-winning dramas. Research on the effects of written fiction suggests that there may be stronger effects for lifetime exposure (Panero et al., 2016). We therefore expected the correlation, if any, between scores on the RMET and the three measures of TV exposure to be positive. Furthermore, we expected the effect to be stronger for award-winning shows and/or characters from award-winning shows.

## Study 1

### Scale Development

Our initial focus was on using character names to gauge familiarity with TV shows. The development of the TCRT was modeled on the development of ART (Acheson et al., 2008; Black et al., 2017; Fong et al., 2013; Mar et al., 2006; Stanovich & West, 1989), designed specifically to assess familiarity with popular TV shows that have aired within the past 10 years. Rather than asking participants to identify the names of TV writers, this test focuses on the names of TV *characters*. Although there are examples of prominent TV creators, such as Shonda Rhimes, Joss Whedon, or J.J. Abrams, these individuals are also frequently involved in other media, such as film. Because previous research has shown that individuals who watch TV frequently form strong bonds with the characters on the screen (Eyal & Cohen, 2006), and because individuals who watch one show may be exposed to promotions for and discussions about other similar shows, we chose to use TV character names to assess exposure instead.

The shows selected for inclusion in the TCRT were chosen from the highest rated (e.g., most viewed, per Nielsen ratings) fictional TV shows airing in the United States in the 2013–2014 season (the year before data collection). The selected shows debuted their first season between 2003 (*NCIS*) and 2013 (*Blacklist*, *Sleepy Hollow*), thus spanning a decade of TV. As of 2015, nearly 80% of the shows had gone into syndication, with reruns airing off-network, and over half were available on streaming services, such as Netflix or Amazon Prime. Roughly a third of the shows had been nomi-

nated for or had won an Emmy in the categories of Best Comedy or Best Drama, and an additional three shows had been similarly recognized for the quality of their lead actor/actress.

A total of 19 shows were selected. For each show, the names of two major characters were generated, with the exception of *Elementary* and *Blue Bloods*, for which one character name was generated apiece. In the case of *Elementary*, this was due to the second major character being Sherlock Holmes, with whom participants would likely be familiar regardless of their exposure to the show. In the case of *Blue Bloods*, the most prominent major characters all shared a single last name and so a single representative was chosen. The decision to include two characters from each show was made to distinguish participants who only recognized the names of iconic characters with a high level of pop cultural visibility (e.g., Sheldon Cooper from *The Big Bang Theory* or Olivia Pope on *Scandal*) from those who had a greater level of exposure to the shows in the question.

The final TCRT has 36 character names and 30 foils. Total scores are computed by subtracting misses from hits. See Table S1 in online supplemental materials for character names.

## Method

**Participants and procedure.** A total of 220 college undergraduates (72.3% female, 95.5% under 21 years old) completed the TCRT and the ART (described in the following text) in lab as part of a larger, Qualtrics-based study. They also answered questions about their book and TV show consumption habits (how many books read on average per month, and how many TV shows followed; see Table 1). All three studies reported in this article were approved by the local institutional review board.

**Author Recognition Test.** The ART (Acheson et al., 2008) is an adaptation of Stanovich and West's (1989) instrument designed to assess exposure to fiction. The ART asks participants to select known authors out of a list of 130 names, half of which are foils. One point is assigned for each hit (correctly identified author), and one point is subtracted for each miss (selected foil).

## Results

**Preliminary analyses.** Participants correctly identified a mean of 4.97 characters ( $SD = 3.58$ ) on the TCRT, or 13.81%, and selected a mean of 0.64 foils ( $SD = 1.35$ ). Three participants had a miss rate greater than 3.0  $SD$  above the mean; one of these, plus two more, were also outliers on ART misses; all five outliers were excluded from subsequent analyses. After excluding outliers, TCRT hits had a mean of 4.89 ( $SD = 3.49$ ); total TCRT scores (hits-misses) had a mean of 4.36 ( $SD = 3.25$ ). Mean ART score (hits-misses) was 9.15 ( $SD = 5.92$ ). Positive skews in both variables were corrected with square root transformations for correlational analyses. There were no gender differences in TCRT,  $t(215) = 0.01$ ,  $p = .994$ , or ART scores,  $t(215) = 0.47$ ,  $p = .637$ , or in TV shows followed,  $t(215) = 1.602$ ,  $p = .111$ .

**Primary analyses.** There was a moderately strong positive correlation between TCRT scores and number of TV shows followed, Spearman's  $\rho$  ( $df = 213$ ) = .39, 95% confidence interval [CI; .27, .50]. As expected, people who reported following more TV shows overall recognized more characters on the TCRT. A very similar relationship was found between ART scores and

Table 1

*Internal Consistency Reliability, Means, and Standard Deviations for Hits on the TV Character Recognition Test (TCRT), Pictorial TCRT (TCRT-P), and TV Show Recognition Test (TSRT); Number of TV Shows Followed; and Self-Report Measures Used*

Scale	$\alpha$	$M$	$SD$
Study 1			
TCRT hits	.75	4.89	3.49
ART hits	.88	9.78	6.13
TV shows followed		3.31	2.75
Books per month		1.97	2.05
Study 2			
TCRT hits	.77	4.88	3.64
ART hits	.88	8.99	6.29
TV shows followed		2.48	0.82
Books per month		1.64	2.18
Parasocial Interactions Scale	.89	68.75	15.43
BFI: Extraversion	.89	3.19	0.84
BFI: Agreeableness	.81	3.76	0.58
BFI: Conscientiousness	.76	3.59	0.53
BFI: Neuroticism	.82	3.05	0.73
BFI: Openness	.75	3.30	0.57
Study 3			
TCRT-P	.90	13.28	8.10
TSRT	.94	27.89	10.33
Parasocial Interactions Scale	.89	72.62	13.73
Social Desirability Scale	.88	31.73	7.26

*Note.* ART = Author Recognition Tests; BFI = Big Five Inventory. Scales used in this research were the ART (Acheson, Wells, & MacDonald, 2008), Parasocial Interactions Scale (Cole & Leets, 1999), BFI (John, Donahue, & Kentle, 1991, 2008), and Crown-Marlowe Social Desirability Scale—Short Form (Reynolds, 1982).

reading habits (number of books read per month), Spearman's  $\rho$  ( $df = 213$ ) = .40, 95% CI [.29, .51]. The relationship between ART and TCRT scores was only slightly weaker,  $r(213) = .34$ , [.23, .47]; those who recognized more authors also tended to recognize more TV characters. See Table 2 for all correlations.

## Discussion

Study 1 provided preliminary evidence that the TCRT's ability to tap TV-viewing behavior was similar to the ART's assessment of reading for pleasure. The correlation between reported TV shows watched and TCRT scores was nearly identical to that between books read for pleasure and ART scores. The next step was to replicate these findings and to investigate the association of TCRT scores with potentially related constructs.

## Study 2

The purpose of Study 2 was to replicate and extend Study 1. Obtaining similar associations between TCRT and ART scores and TV show and reading behavior would validate our initial findings. We also wished to test the relation of TCRT scores with personality and two constructs hypothesized to be related to TV-viewing behavior, theory of mind, as operationalized by the RMET (Baron-Cohen et al., 2001), and the strength of relationship viewers form with fictional characters, operationalized by parasocial interaction (Cole & Leets, 1999; Rubin, Perse, & Powell, 1985). As men-

tioned earlier, experimental studies suggest that viewing award-winning TV dramas (compared with documentaries and control) enhances theory of mind ability (as assessed by the RMET; Black & Barnes, 2015b): As such, it was possible that people who watch more fictional TV shows in general would score higher on the RMET. Similarly, parasocial interaction is related to TV viewing; people who spend more time watching a given show have more opportunities to develop strong emotional attachments to the characters in that show (see Rubin et al., 1985). Moreover, participants who are high in trait parasociability—a disposition toward forming strong relationships with media performers and characters (Schiappa, Allen, & Gregg, 2007)—may become emotionally invested in multiple TV programs, leading them to invest more time and emotion in fictional TV overall. We, therefore, expected positive correlations between TCRT scores and scores on both the RMET and the Parasocial Interactions Scale (described later in the text).

## Method

**Participants and procedure.** As in Study 1, college undergraduates ( $N = 182$ , 76.4% female, 93.2% under 21) completed the study in lab via the Qualtrics platform. After answering questions about their reading and viewing habits, they were presented with the TCRT, ART, and the measures described in the following text. See Table 1 for details.

### Instrumentation.

**Big Five Inventory.** The Big Five Inventory (BFI; John, Donahue, & Kentle, 1991; John, Naumann, & Soto, 2008) assesses the extent to which the traditional big five personality traits (Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness) are dominant in participants based on short phrases. Participants are shown "I am someone who . . ." followed by a list of 44 completions such as "is talkative" and "is relaxed, handles

Table 2

*Correlations Between TV Character Recognition Test (TCRT), Author Recognition Test (ART), Parasocial Interactions Scale, Reading the Mind in the Eyes Test (RMET), and Big Five Inventory (BFI) Scales for Studies 1 and 2*

Variable	TCRT	ART
Study 1		
TV shows followed	.39***	.13
Books read per month	.23**	.40***
Study 2		
TV shows followed	.33***	.12
Books read per month	.15	.41***
Parasocial Interactions Scale	.20**	.18*
RMET	.13	.39***
BFI: Extraversion	-.04	-.13
BFI: Agreeableness	-.08	-.07
BFI: Conscientiousness	.14	.08
BFI: Neuroticism	-.02	.13
BFI: Openness	.13	.23**

*Note.* Pairwise deletion is used for missing data. Spearman's  $\rho$  is presented for correlations between viewing and reading variables (TV shows and books read) and TCRT and ART. The correlation between TCRT and ART scores was  $r = .35$ ,  $p < .001$ , for Study 1, and  $r = .34$ ,  $p < .001$ , for Study 2.

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



stress well.” They indicate their agreement with each phrase on a 5-point Likert scale. John and colleagues (2008) reported very strong correlations with other Big 5 scales. See Table 1 for internal consistency, means, and standard deviations for all scales.

**Reading the Mind in the Eyes Test.** The RMET (Baron-Cohen et al., 2001) was designed to measure theory of mind, or the ability to interpret the emotions and intentions of others, by asking participants to gauge the emotions expressed in the expressions of a series of photographs featuring only the eyes of persons who vary in age, ethnicity, and sex. Participants are shown a picture, and asked to select among four choices the word that best reflects the emotion expressed by the pictured person.

**Parasocial Interactions Scale.** Parasocial interaction refers to the perceived relationship that TV viewers (or possibly fiction readers) feel they have developed with a fictional character or celebrity whom they do not actually know (Cole & Leets, 1999; Eyal & Cohen, 2006). Cole and Leets adapted Rubin et al.’s (1985) instrument in creating the 15-item Parasocial Interactions Scale, which measures the tendency for participants to feel close to favorite TV characters. Items such as “My favorite fictional character makes me feel comfortable, as if I am with friends” are responded to on a 7-point Likert-type scale.

## Results

**Preliminary analyses.** Six cases had miss rates higher than 3.0 *SD* from the mean on the TCRT; five of these and one additional case had excess guessing on the ART: All seven were excluded from subsequent analyses. Mean hit score was 4.88 (*SD* = 3.64), or 13.56%, and mean miss was 0.83 (*SD* = 1.45). Mean total TCRT (hits-misses) was 4.07 (*SD* = 3.50); a slight positive skew was corrected with a square root transformation for correlational analyses. Participants followed a mean of 2.96 TV shows (*SD* = 2.34). Mean ART score was 8.27 (*SD* = 6.14); ART scores had a slight positive skew and were transformed for correlational analyses. All other variables were normally distributed (see Table 1 for descriptive statistics). Females (*M* = 3.13, *SD* = 0.68) scored higher than males (*M* = 2.77, *SD* = 0.85) on the BFI Neuroticism scale,  $t(172) = 2.77, p = .006, d = 0.47$ . There were no other significant gender differences ( $ps > .17$ ).

**Primary analyses.** TCRT scores were moderately and positively correlated with quantity of TV shows followed, Spearman’s  $\rho$  ( $df = 166$ ) = .33, 95% CI [.19, .47], and with ART scores,  $r(166) = .34, [.17, .48]$ . The relationship between ART scores and reported number of books read was similar to that between TCRT and TV viewing, Spearman’s  $\rho$  ( $df = 172$ ) = .41, [.26, .53]. The correlation between TCRT and ART scores was very similar to that found in Study 1,  $r(166) = .34, [.17, .50]$ . TCRT scores were not related to performance on the RMET,  $r(166) = .13, [-.06, .31]$ , whereas ART scores were,  $r(166) = .40, [.26, .53]$ . TCRT scores had a moderately weak positive correlation with parasocial interactions,  $r(165) = .20, [.06, .34]$ , similar to that between ART scores and parasocial interactions,  $r(172) = .18, [.05, .31]$ . TCRT scores were not related to any of the BFI scales. See Table 2 for correlations between the TCRT, ART, and all variables.

## Discussion

Overall, results from Study 2 validated the findings in Study 1. TCRT scores had a moderate positive correlation with reported TV

shows followed, and weaker correlation with books read for pleasure, whereas ART scores had moderately strong positive correlations with books read and weak, nonsignificant correlations with TV shows followed. The association between TCRT and ART scores was the same across the two studies ( $r_s = .34, .35$ ). Both the TCRT and the ART were positively correlated with the Parasocial Interactions Scale. Interestingly, although ART scores were positively related to RMET scores and Openness to Experience, TCRT scores were not, nor were they related to any other personality trait. Given that one-off exposure to award-winning dramas has been shown to improve performance on the RMET (Black & Barnes, 2015b), the lack of association with TCRT scores was intriguing. It could be that there was in fact no effect, or it could be that our measurement of TV exposure did not capture the type of TV-viewing behavior that would relate to theory of mind. In Study 3, we addressed both issues by testing different recognition formats and exploring the difference between award-winning and non-award-winning shows.

## Study 3

The TCRT used in Studies 1 and 2 provided normally distributed data, but mean character recognition was low, and the names had been chosen from shows broadcast on TV. People who stream TV shows on a regular basis might not have been exposed to those characters; not only do companies such as Netflix and Amazon Prime provide access to traditional shows, they also feature streaming originals, including shows that are *only* available through their services. A further issue with the character name test is that many people may know the characters only by their first (or last) names; even if the researcher’s focus is purely network TV, participants who do not recognize the full name may introduce measurement error. We therefore modified the TCRT in two ways: First, we added 25 characters from shows that were either streaming originals or popular on streaming. Second, we added pictures of the characters, such that participants would have to recognize the character name, aided by the image, as well as the actor. Foils were paired with headshots similar in style and framing to those of the TV characters. The final pictorial TCRT (TCRT-P) had 61 TV characters and 30 foils.

Adding images to the TCRT along with characters from streamed shows was predicted to improve the instrument, but it also made it more difficult to administer. Further, it is possible that asking participants to simply identify the names of the *shows* would be sufficient. Past researchers have used only show names as a proxy for TV exposure; for example, Mar and colleagues (2009) used parents’ recognition of children’s TV shows and movies as an index of exposure. We therefore designed the TSRT to compare with the TCRT. The TSRT includes all of the shows represented on the TCRT, plus 19 similar shows, for a total of 51 hits. Twenty-four foil titles were generated by a professional writer with experience writing TV pilots. See Tables S1 (names from Studies 1 and 2 with associated shows) and S2 (new names/shows) in the online supplemental materials for TCRT and TSRT names and hit rates.

## Award-Winning Versus Non-Award-Winning Shows

We classified shows as award-winning if they met the following criteria: nomination for best comedy or drama *or* nomination for

two or more of the categories of directing, writing, or best leading actor or actress. For the TSRT, there were 17 award-winning shows. All but four of these had characters on the TCRT, for a total of 23 characters in the award-winning category.

## Method

A total of 226 adults (45.6% female, mean age = 34) were recruited on Amazon.com's Mechanical Turk (MTurk) and paid \$2.40 to complete an online survey offered on Qualtrics. Of these, five were discarded for excessive guessing (more than 3 *SD* above the mean foils selection) on the TCRT-P, one was discarded for excessive guessing on the TSRT, and one was discarded for excessive guessing on both measures. Participants completed first the TSRT and demographic questions, followed by the RMET, and two questionnaires, the Parasocial Interactions Scale described previously, and a measure of socially desirable responding, detailed in the following text. They also indicated how they watched TV shows (TV, streaming, DVD; see Table S3 in the online supplemental materials for details). Finally, participants completed the TCRT-P. See Table 1 for means and internal consistency reliability for all measures.

Socially desirable response bias was assessed with the Crown-Marlowe Social Desirability Scale Short Form (CMSDS-SF; Reynolds, 1982). Because self-reported reading and TV watching behavior may be over- and underreported respectively to meet social norms, testing for potential correlations between TCRT scores and social desirability bias would lend credence to subsequent analyses: If scores were related to socially desirable responding, then scores on the CMSDS-SF could be included as covariate. The CMSDS-SF is a 13-item questionnaire scored in the direction of greater socially desirable responding. In this study, a 4-point scale (*strongly disagree* to *strongly agree*) was used; sample items include "No matter who I'm talking to, I'm always a good listener," and "I have never deliberately said something that hurt someone's feelings."

## Results

**Preliminary analyses.** Participants reported following a mean of 4.57 (*SD* = 3.21) TV shows. Not including one outlier that had selected 25 foils, participants selected a mean of 0.56 (*SD* = 1.36) foils on the TCRT-P. After discarding this and four more outliers for excess guessing, participants identified a mean of 13.69 characters (*SD* = 8.24, range 0 to 42), or 22.44%, out of the full 61-character TCRT-P. Of the 36 names used in Studies 1 and 2, participants identified a mean of 8.70 names (*SD* = 6.09), or 24.17%. After subtracting misses, mean TCRT-P was 13.28 (*SD* = 8.10); a slight positive skew was corrected with a square root transformation for correlational analyses. On the TSRT, participants selected a mean of 0.89 foils (*SD* = 1.01). Not counting two outliers, mean TV shows selected was 28.76 (*SD* = 10.70, range 1 to 50), or 56.36%, out of 51 hits. After subtracting misses, mean TSRT was 27.89 (10.33). TSRT scores were normally distributed. See Table S3 in the online supplemental materials for means on all scales by viewing habits.

To compare mean recognition of award-winning versus non-award-winning shows and characters from them, percentages were used. Total scores were used for correlational analyses; the non-

award-winning character variable was positively skewed; a square root transformation achieved normal distribution. RMET scores were negatively skewed; a reflected square root transformation corrected the variable for correlational analyses. Age was positively skewed; a logarithmic transformation was used to correct the skew. All other variables were normally distributed.

**Gender differences and age.** Women (*M* = 15.49, *SD* = 8.20) recognized more characters than men (*M* = 12.14, *SD* = 7.88) did,  $t(217) = 3.06$ ,  $p = .003$ ,  $d = 0.41$ . Women (*M* = 30.23, *SD* = 10.35) also tended to recognize more TV shows than men (*M* = 27.54, *SD* = 10.92), but the difference was not statistically significant,  $t(221) = 1.89$ ,  $p = .062$ ,  $d = 0.25$ . Age was not related to TCRT-P or TSRT scores, parasocial interactions, or social desirability, although it was related to the number of TV shows followed (Spearman's  $\rho = .24$ ,  $p < .001$ ) and RMET scores,  $r = .19$ ,  $p = .005$ .

**Socially desirable response bias.** Social desirability scores were not related to reported TV viewing (Spearman's  $\rho = .03$ ,  $p = .706$ ), but they were related to reported book reading ( $\rho = .16$ ,  $p = .013$ ): The more books participants reported reading each month, the greater their scores on the Social Desirability Scale. Social desirability was not related to TCRT-P scores,  $r = -.10$ ,  $p = .125$ , but it was negatively related to TSRT scores,  $r = -.16$ ,  $p = .019$ ; people who recognized more shows gave fewer socially desirable responses.

### Primary analyses.

**Comparison of the TCRT-P and the TSRT.** The two recognition tests were strongly correlated,  $r(216) = .65$ , 95% CI [.53, .73]. To compare the two tests, percentage scores were calculated by dividing the total scores (corrected for guessing) by the number of hits on each test. Participants correctly identified many more TV shows (56.31%) than characters (22.48%),  $t(217) = 30.33$ ,  $p < .001$ ,  $d = 1.91$ . The same pattern held for both males and females: On average, participants recognized over 50% of the TV shows, and approximately one fifth of the characters. Both TSRT (Spearman's  $\rho = .55$ , [.47, .64]) and TCRT-P ( $\rho = .51$ , [.41, .60]) scores were related to reported TV viewing behavior (number of shows followed).

People who recognized more TV shows did better on the RMET,  $r = .16$ , 95% CI [.02, .31]. This effect was driven by men,  $r(119) = .22$ ,  $p = .015$ . For women, TV show recognition was not related to RMET scores,  $r(100) = .09$ ,  $p = .391$ . The correlation between RMET and TCRT-P score was also positive ( $r = .13$ , [-.02, .26]), but was not statistically significant ( $p = .061$ ); this relation was similar for women and men. Neither TCRT-P nor TSRT scores were related to parasocial interactions,  $ps > .39$ . See Table 3 for all correlations.

**Pictorial TCRT.** Participants recognized more characters from award-winning shows (*M* = 28.6%, *SD* = 16.7%) than from nonaward winners (*M* = 18.8%, *SD* = 14.0%),  $t(217) = 10.41$ ,  $p < .001$ ,  $d = 0.63$ . However, people who recognized more characters from award-winning shows also recognized more characters from non-award-winning shows,  $r = .61$ ,  $p < .001$ . Scores on the RMET and Parasocial Interactions Scale were not related to recognition of characters from award-winning ( $rs = .07$ ;  $ps > .30$ ) or non-award-winning ( $rs = .10$ , .08;  $ps > .14$ ) shows. There were no gender differences.

**TV Show Recognition Test.** Similarly, participants recognized more titles from award-winning show (*M* = 73.0%, *SD* = 20.3%)

Table 3

*Zero-Order Correlations Between Pictorial TV Character Recognition Test (TCRT-P) and TV Show Recognition Test (TSRT; Totals and Broken Down by Award- vs. Non-Award-Winning Shows), TV Shows Followed, Reading the Mind in the Eyes Test (RMET), Parasocial Interactions Scale, and Social Desirability Scale*

	2	3	4	5	6	7	8	9	10
1. TCRT-P	.87***	.90***	.65***	.58***	.65***	.51***	.12	.06	-.11
2. Award-winning shows	—	.61***	.60***	.58***	.59***	.52***	.07	.07	-.11
3. Non-award-winning shows		—	.52***	.43***	.55***	.39***	.10	.08	-.07
4. TSRT			—	.92***	.98***	.55***	.15*	-.04	-.16*
5. Award-winning shows				—	.84***	.50***	.18**	-.04	-.13
6. Non-award-winning shows					—	.54***	.14*	-.03	-.15*
7. TV shows followed						—	.12	.02	.02
8. RMET							—	-.13	.03
9. Parasocial Interactions Scale								—	-.01
10. Social Desirability Scale									—

Note.  $N = 218$ . Spearman's  $\rho$  is presented for correlations with TV shows followed.

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

than from nonaward winners ( $M = 48.0\%$ ,  $SD = 22.8\%$ ),  $t(217) = 28.00$ ,  $p < .001$ ,  $d = 1.15$ . The correlation between award-winning and non-award-winning shows recognized was very strong,  $r = .82$ ,  $p < .001$ . Scores on the Parasocial Interactions Scale were not related to recognition of either award-winning or non-award-winning shows ( $r_s = -.04$ ,  $-.03$ ,  $p_s > .50$ ). Both were positively correlated with RMET scores, with the relationship being stronger for award-winning ( $r = .18$ ,  $p = .007$  vs.  $r = .14$ ,  $p = .047$ ). However, this relationship was driven by men: For men, the correlation between award-winning shows and RMET scores was  $r(114) = .24$ ,  $p = .011$ , which was stronger ( $z = -1.73$ ,  $p_{\text{one-tailed}} = .042$ ) than the correlation between non-award-winning shows and RMET scores,  $r = .18$ ,  $p = .057$ . For women, neither correlation was significant ( $r_s < .12$ ,  $p_s > .24$ ).

## Discussion

The purpose of Study 3 was twofold: We wanted to test two different formats of TV recognition tests, and we wanted to explore potential differences in exposure to award-winning versus non-award-winning TV. The pictorial TCRT (pictures of characters) and the TSRT (names of shows) were strongly related to reported TV-viewing behavior (number of shows followed). Neither was associated with strength of parasocial interaction, somewhat surprisingly given the positive correlation between (nonpictorial) TCRT and the Parasocial Interactions Scale found in Study 2. It is unclear if this is a result of differences in parasociability in undergraduate and MTurk samples; previous research has shown inconsistent correlations between strength of parasocial interaction and age (Schiappa et al., 2007). Future research is needed to explore these differences. Importantly, both the TCRT-C and the TSRT seem to function similarly. Socially desirable responding does not appear to be an issue for reporting TV-viewing behavior, although there was a weak negative correlation with TSRT scores. The only meaningful difference between the two measures was for associations with theory of mind: Only the TSRT was related to RMET scores, and the association was strongest for award-winning shows and men. Given the overall similarity, we believe both of these additional measures can be of use to researchers studying TV exposure, depending on whether the theoretical

grounding of the study and other measures used are more geared toward TV characters or TV shows as a whole.

## General Discussion

The purpose of this study was to provide measures of exposure to TV for use in research where previous experience may confound the effect of the experimental manipulation, and in correlational studies investigating relationships between TV-viewing habits and relevant behavior or traits. We present three such measures, the TCRT in two forms (only character names and a pictorial version with images), and the TSRT. We used the RMET as a comparison measure both because experimental research has shown that exposure to award-winning dramas can improve performance on the RMET (Black & Barnes, 2015b), and because the RMET has been used frequently in research on the effects of fiction (Black & Barnes, 2015a; Kidd & Castano, 2013; Mar et al., 2006, 2009; Panero et al., 2016).

We modeled the TCRT and TSRT on a widely used instrument designed to assess lifetime exposure to literature, the ART (Stanovich & West, 1989). Both measures rely on name recognition that could come from both direct experience (watching the show or reading the book) and experience with related content, such as a specific genre (e.g., science fiction or romance) or programming (although people might not watch a show on our list, they might be familiar with it if scheduled immediately before or after another favorite show, or if the two shows are cross-promoted by the viewing algorithms on a given streaming platform). Across two studies, the relationship between the ART and the number of books participants reported reading per month and the relationship between the TCRT and the number of TV shows participants reported following was very similar. Although the purpose of using checklists to assess exposure rather than self-report is partly to avoid socially desirable responding (e.g., participants may feel compelled to report that they read more books than they actually do or that they spend less time watching TV), it is important that participants' accounts of their preferences have some relationship with parallel measures of exposure (see Rain & Mar, 2014). That the two measures captured roughly equal amounts of variation in self-reported habits suggests that one can measure TV exposure

using instruments such as the TCRT, though our results also suggest that using a recognition test may be less necessary in assessing general TV exposure; in Study 3, we measured socially desirable response bias and found that although people who gave more socially desirable responses reported reading more books, they did not report following fewer TV shows.

Interestingly, in Study 1, whereas scores on the TCRT had moderately weak positive correlations with the number of books participants reported reading per month (although the effect was in the same direction in Study 2, it did not reach statistical significance [ $p = .057$ ]), ART scores were not related to the number of TV shows an individual reported currently following. In other words, although individuals who scored higher on our measure of TV exposure were also likely to report reading more books, lifetime exposure to written fiction was not predictive of current TV-viewing habits. One possible reason for this pattern of results is that following a TV show may require a greater level of commitment—both in terms of time and long-term emotion—than reading a book. It is thus possible that individuals who are more willing to make this commitment may be highly drawn to fictional stories across media, whereas individuals who score high on the ART may be more focused on the written word.

Another possible explanation for this distinction is that although the ART is a measure of lifetime print exposure, the TCRTs and TSRT may be more accurately seen as a measure of cumulative TV viewing over the past few years. In contrast to the ART, which features authors published throughout the 20th century (some in fact began before then, and others continue today), the characters and shows on the TV recognition tests were taken from shows popular over the decade before testing. This aspect of the TCRT may make it a better measure of cumulative TV exposure over a period of years, rather than lifetime exposure per se; similarly, the TCRT/TSRT may specifically capture voluntary, recreational viewing, whereas the ART may also capture exposure to literature in the classroom. Researchers wishing to capture exposure to TV over distinct periods of time can easily adapt the method by choosing characters and shows pertaining to that specific era. Alternatively, they could target specific TV genres by including enough names in the targets along with other shows and foils (see Black et al., 2017, for a similar treatment of book genres).

In Study 3, we chose to target award-winning versus non-award-winning shows, because past research suggests that quality fiction may have a stronger relation to scores on the RMET (Black & Barnes, 2015b; Kidd & Castano, 2013). It is important to note, however, that not only may the effect of literary fiction be tenuous and difficult to replicate (see Panero et al., 2016), but that the effect of past exposure (as assessed by the ART of Acheson et al., 2008) was the strongest predictor of RMET scores in both the experimental studies mentioned earlier. Kidd and Castano (2017) divided the Acheson et al.'s ART into literary and genre fiction and compared their relations with RMET scores: Although genre and literary fiction had the same zero-order correlation with RMET scores, when entered together in regression analyses, only the effect of literary fiction was significant. We used similar analyses in Study 3 to show that familiarity with award-winning shows was uniquely related to RMET scores. Curiously, we did not find the same effect for names of characters from award-winning shows. This may be because there were more shows represented in the TVST than the TVCT-P, thus providing more

opportunities for name recognition. On the other hand, character name recognition should be a better indicator of real familiarity. It could be that award-winning show recognition is also a sign of general erudition; award-winning shows are reviewed in literary magazines that may be read by people who never actually watch the shows (and are therefore unlikely to recognize the characters). It has been suggested that performance of the RMET also reflects education due to its vocabulary demands (Panero et al., 2016).

It may also be that third variables explain the relationship between award-winning show recognition and RMET scores; however, these results are in line with past research that shows a unique effect of award-winning TV (Black & Barnes, 2015b). That said, the current research also suggests that those who are familiar with award-winning shows may be more familiar with non-award-winning shows as well; the correlation between award-winning versus not was very strong for both character and show name recognition. Moreover, the current findings also suggest that some people may be more drawn to fictional stories no matter the form; the relationship between TCRT scores and ART scores was consistent across studies ( $r_s = .34$  and  $.35$ ).

However, people who are drawn to both novels and TV shows may not share many other characteristics. TCRT scores were not significantly correlated with any of the Big 5 personality factors; ART scores were correlated with Openness ( $r = .23$ ) but not with any other traits, which is consistent with the results reported by Mar et al. (2009). Overall TV exposure was not related to theory of mind as operationalized by the scores on the RMET. Research on the association between lifetime exposure to fiction and RMET scores provides robust evidence of a positive correlation (Mumper & Gerrig, 2017; Panero et al., 2016); the effect we found in Study 2 ( $r = .39$ ) is in line with past research. Here, we found no relation between TV character recognition and RMET scores, and the positive correlation with the TSRT was driven by recognition of award-winning shows (not entirely surprising given experimental evidence of the immediate effects of viewing award-winning dramas; Black & Barnes, 2015b) and was true only for males. The sex difference could be important: Males tend to perform worse on the RMET and measures of empathy in general (Baron-Cohen, 2010). Given that the RMET is an easy test for neurotypical adults in general (Black, 2018), it could be that few are seeing a ceiling effect for females, and better measurement of theory of mind would reveal a relation with TV viewing. Alternatively, it could be that TV exposure is associated with greater theory of mind skills for males in particular, or, more generally, populations that tend to perform worse on theory of mind tasks. Recent research that has extended work on lifetime written fiction exposure and theory of mind to a Latin American population found a similar pattern of results, with the correlation only significant for male participants (Tabullo, Jiménez, & García, 2018).

Given that a great deal of previous research has explored the relationship between performance on author recognitions tests and a variety of dependent measures such as empathy (Mar et al., 2006), verbal ability (Mar & Rain, 2015), gender role attitudes (Fong et al., 2013), and personality (Mar et al., 2009), there is a need for a measure that will allow researchers to explore the degree to which these results may generalize across media. Both the TCRT and the TSRT assess exposure using the same method as the various author recognition tasks, providing consistency of measurement. Although social desirability—an important feature



of recognition tasks is that they are presumed to mitigate response bias—was not related to reported TV viewing behavior (Study 3), such questions (“How many TV shows do you follow/books do you read?”) refer to quantity, not content. Targeting particular types of shows might make social desirability an issue. One of the advantages of the method outlined here is that it provides a template for researchers to generate further TV recognition tests—for example, recognition tests that measure exposure to reality TV or fantasy, or shows with female protagonists or racially diverse casts.

## References

- Acheson, D. J., Wells, J. B., & MacDonald, M. C. (2008). New and updated tests of print exposure and reading abilities in college students. *Behavior Research Methods*, 40, 278–289. <http://dx.doi.org/10.3758/BRM.40.1.278>
- Baron-Cohen, S. (2010). Empathizing, systemizing, and the extreme male brain theory of autism. In I. Savic (Ed.), *Sex differences in the human brain, their underpinnings and implications* (pp. 167–175). Amsterdam, the Netherlands: Elsevier. <http://dx.doi.org/10.1016/B978-0-444-53630-3.00011-7>
- Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The “Reading the Mind in the Eyes” Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry*, 42, 241–251. <http://dx.doi.org/10.1111/1469-7610.00715>
- Black, J. E. (2018). An IRT analysis of the Reading the Mind in the Eyes test. *Journal of Personality Assessment*. Advance online publication. <http://dx.doi.org/10.1080/00223891.2018.1447946>
- Black, J. E., & Barnes, J. L. (2015a). The effects of reading material on social and non-social cognition. *Poetics*, 52, 32–43. <http://dx.doi.org/10.1016/j.poetic.2015.07.001>
- Black, J. E., & Barnes, J. L. (2015b). Fiction and social cognition: The effect of viewing award-winning television dramas on theory of mind. *Psychology of Aesthetics, Creativity, and the Arts*, 9, 423–429. <http://dx.doi.org/10.1037/aca0000031>
- Black, J. E., Capps, S. C., & Barnes, J. L. (2017). Fiction, genre exposure, and moral reality. *Psychology of Aesthetics, Creativity, and the Arts*, 12, 328–340. <http://dx.doi.org/10.1037/aca0000116>
- Bortolussi, M., & Dixon, P. (2003). *Psychonarratology: Foundations for the empirical study of literary response*. New York, NY: Cambridge University Press.
- Christakis, D. A., Zimmerman, F. J., DiGiuseppe, D. L., & McCarty, C. A. (2004). Early television exposure and subsequent attentional problems in children. *Pediatrics*, 113, 708–713. <http://dx.doi.org/10.1542/peds.113.4.708>
- Cipielewski, J., & Stanovich, K. E. (1992). Predicting growth in reading ability from children’s exposure to print. *Journal of Experimental Child Psychology*, 54, 74–89. [http://dx.doi.org/10.1016/0022-0965\(92\)90018-2](http://dx.doi.org/10.1016/0022-0965(92)90018-2)
- Cole, T., & Leets, L. (1999). Attachment styles and intimate television viewing: Insecurely forming relationships in a parasocial way. *Journal of Social and Personal Relationships*, 16, 495–511. <http://dx.doi.org/10.1177/0265407599164005>
- Dilliplane, S., Goldman, S. K., & Mutz, D. C. (2013). Televised exposure to eolitics: New measures for a fragmented media environment. *American Journal of Political Science*, 57, 236–248. <http://dx.doi.org/10.1111/j.1540-5907.2012.00600.x>
- Djikic, M., Oatley, K., & Moldoveanu, M. C. (2013a). Opening the closed mind: The effect of exposure to literature on the need for closure. *Creativity Research Journal*, 25, 149–154. <http://dx.doi.org/10.1080/10400419.2013.783735>
- Djikic, M., Oatley, K., & Moldoveanu, M. C. (2013b). Reading other minds: Effects of literature on empathy. *Scientific Study of Literature*, 3, 28–47. <http://dx.doi.org/10.1075/ssol.3.1.06dji>
- Eyal, K., & Cohen, J. (2006). When good *Friends* say goodbye: A parasocial breakup study. *Journal of Broadcasting and Electronic Media*, 50, 502–523. [http://dx.doi.org/10.1207/s15506878jobem5003\\_9](http://dx.doi.org/10.1207/s15506878jobem5003_9)
- Fong, K., Mullin, J. B., & Mar, R. A. (2013). What you read matters: The role of fiction genre in predicting interpersonal sensitivity. *Psychology of Aesthetics, Creativity, and the Arts*, 7, 370–376. <http://dx.doi.org/10.1037/a0034084>
- Hakemulder, J. (2000). *The moral laboratory: Experiments examining the effects of reading literature on social perception and moral self-concept*. Amsterdam, the Netherlands: John Benjamins Publishing Company.
- Huesmann, L. R., Moise-Titus, J., Podolski, C.-L., & Eron, L. D. (2003). Longitudinal relations between children’s exposure to TV violence and their aggressive and violent behavior in young adulthood: 1977–1992. *Developmental Psychology*, 39, 201–221. <http://dx.doi.org/10.1037/0012-1649.39.2.201>
- Hummer, T. A., Kronenberger, W. G., Wang, Y., Anderson, C. C., & Mathews, V. P. (2014). Association of television violence exposure with executive functioning and white matter volume in young adult males. *Brain and Cognition*, 88, 26–34. <http://dx.doi.org/10.1016/j.bandc.2014.04.010>
- John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The Big Five Inventory—Versions 4a and 54*. Berkeley: University of California, Institute of Personality and Social Research.
- John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (pp. 114–158). New York, NY: Guilford Press.
- Kidd, D. C., & Castano, E. (2013). Reading literary fiction improves theory of mind. *Science*, 342, 377–380. <http://dx.doi.org/10.1126/science.1239918>
- Kidd, D. C., & Castano, E. (2017). Different stories: How levels of familiarity with literary and genre fiction relate to mentalizing. *Psychology of Aesthetics, Creativity, and the Arts*, 11, 474–486. <http://dx.doi.org/10.1037/aca0000069>
- Koopman, E. M. E., & Hakemulder, F. (2015). Effects of literature on empathy and self-reflection: A theoretical-empirical framework. *Journal of Literary Theory*, 9, 79–111. <http://dx.doi.org/10.1515/jlt-2015-0005>
- Mar, R. A. (2018). Evaluating whether stories can promote social cognition: Introducing the social processes and content entrained by narrative (SPaCEN) framework. *Discourse Processes*, 55, 454–479. <http://dx.doi.org/10.1080/0163853X.2018.1448209>
- Mar, R. A., & Oatley, K. (2008). The function of fiction is the abstraction and simulation of social experience. *Perspectives on Psychological Science*, 3, 173–192. <http://dx.doi.org/10.1111/j.1745-6924.2008.00073.x>
- Mar, R. A., Oatley, K., Hirsh, J., de la Paz, J., & Peterson, J. (2006). Bookworms versus nerds: Exposure to fiction versus non-fiction, divergent associations with social ability, and the simulation of fictional social worlds. *Journal of Research in Personality*, 40, 694–712. <http://dx.doi.org/10.1016/j.jrp.2005.08.002>
- Mar, R. A., Oatley, K., & Peterson, J. B. (2009). Exploring the link between reading fiction and empathy: Ruling out individual differences and examining outcomes. *Communications*, 34, 407–428. <http://dx.doi.org/10.1515/COMM.2009.025>
- Mar, R. A., & Rain, M. (2015). Narrative fiction and expository nonfiction differentially predict verbal ability. *Scientific Studies of Reading*, 19, 419–433. <http://dx.doi.org/10.1080/10888438.2015.1069296>
- Mar, R. A., Tackett, J. L., & Moore, C. (2010). Exposure to media and theory-of-mind development in preschoolers. *Cognitive Development*, 25, 69–78. <http://dx.doi.org/10.1016/j.cogdev.2009.11.002>

- Mumper, M. L., & Gerrig, R. J. (2017). Leisure reading and social cognition: A meta-analysis. *Psychology of Aesthetics, Creativity, and the Arts*, 11, 109–120. <http://dx.doi.org/10.1037/aca0000089>
- Nathanson, A. I., Aladé, F., Sharp, M. L., Rasmussen, E. E., & Christy, K. (2014). The relation between television exposure and executive function among preschoolers. *Developmental Psychology*, 50, 1497–1506. <http://dx.doi.org/10.1037/a0035714>
- Oatley, K. (2016). Fiction: Simulation of social worlds. *Trends in Cognitive Sciences*, 20, 618–628. <http://dx.doi.org/10.1016/j.tics.2016.06.002>
- Panero, M. E., Weisberg, D. S., Black, J., Goldstein, T. R., Barnes, J. L., Brownell, H., & Winner, E. (2016). Does reading a single passage of literary fiction really improve theory of mind? An attempt at replication. *Journal of Personality and Social Psychology*, 111, e46–e54. <http://dx.doi.org/10.1037/pspa0000064>
- Potter, W. J., & Chang, I. C. (1990). Television exposure measures and the cultivation hypothesis. *Journal of Broadcasting and Electronic Media*, 34, 313–333. <http://dx.doi.org/10.1080/08838159009386745>
- Rain, M., & Mar, R. A. (2014). Measuring reading behavior: Examining the predictive validity of print-exposure checklists. *Empirical Studies of the Arts*, 32, 93–108.
- Reynolds, W. M. (1982). Development of reliable and valid short forms of the Marlow–Crowne Social Desirability Scale. *Journal of Clinical Psychology*, 38, 119–125. [http://dx.doi.org/10.1002/1097-4679\(198201\)38:1<119::AID-JCLP2270380118>3.0.CO;2-I](http://dx.doi.org/10.1002/1097-4679(198201)38:1<119::AID-JCLP2270380118>3.0.CO;2-I)
- Rubin, A., Perse, E., & Powell, R. (1985). Loneliness, parasocial interaction, and local television news viewing. *Human Communication Research*, 12, 155–180. <http://dx.doi.org/10.1111/j.1468-2958.1985.tb00071.x>
- Samur, D., Tops, M., & Koole, S. L. (2018). Does a single session of reading literary fiction prime enhanced mentalising performance? Four replication experiments of Kidd and Castano (2013). *Cognition and Emotion*, 32, 130–144. <http://dx.doi.org/10.1080/02699931.2017.1279591>
- Schiappa, E., Allen, M., & Gregg, P. B. (2007). Parasocial relationships and television: A meta-analysis of the effects. In R. W. Preiss, B. M. Gayle, N. Burrell, M. Allen, & J. Bryant (Eds.), *Mass media effects research: Advances through meta-analysis* (pp. 301–314). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, and Computers*, 31, 137–149. <http://dx.doi.org/10.3758/BF03207704>
- Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly*, 24, 402–433. <http://dx.doi.org/10.2307/747605>
- Stanovich, K. E., Cunningham, A. E., & West, R. F. (1998). Literacy experiences and the shaping of cognition. In S. G. Paris & H. M. Wellman (Eds.), *Global prospects for education: Development, culture, and schooling* (pp. 253–288). Washington, DC: American Psychological Association. <http://dx.doi.org/10.1037/10294-009>
- Tabullo, A. J., Jiménez, V. A. N., & García, C. S. (2018). Associations between fiction reading, trait empathy and theory of mind ability. *International Journal of Psychology and Psychological Therapy*, 18, 357–370.
- van Lissa, C. J., Caracciolo, M., van Duuren, T., & van Leuven, B. (2016). Difficult empathy: The effect of narrative perspective on readers' engagement with a first-person narrator. *DIEGESIS*, 5. Retrieved from <https://www.diegesis.uni-wuppertal.de/index.php/diegesis/article/view/211/293>
- Webster, J. G., & Wakshlag, J. (2013). Measuring exposure to television. In D. Zillmann & J. Bryant (Eds.), *Selective exposure to communication* (pp. 35–63). New York, NY: Routledge.
- Wonneberger, A., & Irazoqi, M. (2016). Explaining response errors of self-reported frequency and duration of TV exposure through individual and contextual factors. *Journalism and Mass Communication Quarterly*, 94, 1–23.
- Zunshine, L. (2006). *Why we read fiction: Theory of mind and the novel*. Columbus: Ohio State University Press.

Received April 24, 2019

Revision received June 12, 2019

Accepted June 13, 2019 ■

### E-Mail Notification of Your Latest Issue Online!

Would you like to know when the next issue of your favorite APA journal will be available online? This service is now available to you. Sign up at <https://my.apa.org/portal/alerts/> and you will be notified by e-mail when issues of interest to you become available!