ARTICLE

Check for updates

Art films foster theory of mind

Emanuele Castano [1,2 ✉]

Research shows that reading literary but not popular fiction enhances the Theory of Mind (ToM). This article builds on the symmetry between literary theory and film theory and investigates whether exposure to art films, but not Hollywood films, enhances ToM. Participants ($N = 232$) were randomly assigned to view either art or Hollywood films and then answered questions about the film and its characters before completing two separate measures of ToM (the Read the Mind in the Eyes Test and the Moral Judgement Task). Results showed that art film viewers scored higher on both ToM measures and that the effect was sequentially mediated by perception of complexity and predictability of the characters. The findings are discussed in the context of the emerging literature on the impact of fiction on social cognition.

[1] University of Trento, Trento, Italy. [2] National Research Council, Rome, Italy. ✉email: emanuele.castano@unitn.it

Introduction

Just 100 years ago Soviet filmmaker Kuleshov showed, to different audiences, footage in which a character's neutral facial expression was followed by the image of a plate of soup, a dead girl in a coffin, or a woman on a sofa. Subsequently asked about the mental state of the character, spectators from the three audiences gave different answers, inferring hunger, grief, or desire, respectively. This is known as the Kuleshov effect: a spectator's inferences about the mental states of a character on the screen depend on the broader context in which the character appears. The Kuleshov effect elucidates the art of feature-length film-making, through which spectators' perceptual and cognitive processes, as well as sociocultural knowledge, are leveraged to create the cinematic experience. In this article, I propose that characteristics of the film influence not only what spectators infer about the characters' mental states, but also the spectators' accuracy in making inferences about the mental states of others, outside of the fictional world of film.

Theory of mind

Inferring the mental states of others is known as mentalizing (Frith, 1989) or Theory of Mind (ToM; Premack and Woodruff, 1978). ToM is not a unitary process, but rather a system that develops both as a set of abstract rules about the functioning of people's mind (Gopnik and Wellman, 1992) and through the simulation of their experiences (Gordon, 1986)—aided at least in part by mirror neurons (Gallese and Goldman, 1998). In its most basic form, which is already present in the first years of life (Heyes and Frith, 2014), ToM consists in recognizing that others have a mind that differs from our own (Premack and Woodruff, 1978). In adulthood, ToM refers to the capacity to accurately representing other people's thoughts, intentions, emotions, beliefs, and desires. The extensive literature on ToM has further differentiated between its affective and cognitive components (Decety and Jackson, 2016; Shamay-Tsoory & Aharon-Peretz, 2007), leading to some overlap between ToM and the concept of empathy. However, while empathy is typically considered affect-sharing, affective ToM is the capacity, perhaps also aided by affect-sharing, to decode mental states in others, and the two are supported by different neurological networks (Kanske et al., 2015).

The multidimensionality of the ToM construct is further highlighted by the development of different classes of tasks to measure it. Some are socio-perceptual in nature, capturing, for instance, variability in the capacity to infer mental states from faces (Baron-Cohen et al., 2001). Other, socio-cognitive tasks, typically use the brief written description of scenarios and tap differences in the ability of the readers to represent the varying degree of embedded mental states (he thinks, that she thinks that he thinks….) of the actors, or to infer and consider the actors' intentions when judging, for instance, the admissibility of their actions (Young et al., 2007). The literature is not yet clear as to whether individual differences on these measures are capturing actual variability in ToM capacity, reflect an interest in reading other people's minds, or are the consequence of heterogeneity in the cognitive apparatus that supports ToM (Apperly, 2012). These are important questions, to which I shall return in the discussion. For now, we concern ourselves with differences in performance in well-established measures of ToM.

Substantial heterogeneity exists in the performance of ToM tasks in adult populations. Research has shown that ToM is enhanced by learning a second language (Rubio-Fernández and Glucksberg, 2012), acting (Goldstein and Winner, 2012), meditation practices (Mascaro et al., 2013), and, particularly, reading fiction (Mar and Oatley, 2008; Oatley, 2016). Exposure to fiction

predicts performance on ToM tasks, even after controlling for variables such as age, gender, experience with the English language, intelligence, transportation (in the narrative), and personality traits (Mar et al., 2006; Mar et al., 2009). It also matters what kind of fiction one reads. A series of correlational studies have shown that it is specifically exposure to literary fiction that predicts performance on ToM tasks, while exposure to popular fiction does not (Castano et al., 2020; Kidd and Castano, 2017b). Experimental studies also support this conjecture. In a series of five experiments participants were randomly assigned to read excerpts of novels or entire short stories, that were classified as either literary or popular. Results showed that those participants who had read literary stories subsequently performed better on a variety of tasks assessing their ToM (Kidd and Castano, 2013). Several other experiments replicated this finding (Black and Barnes, 2015a; Kidd and Castano, 2013, 2019; Kidd et al., 2016; Pino and Mazza, 2016; van Kuijk et al., 2018). Those that did not (e.g., Panero et al., 2016) suffer from methodological flaws that make them hard to interpret (Kidd and Castano, 2017b).

It thus seems that the effect of fiction on ToM is specific to literary fiction. What is this due to? What are the idiosyncratic features of literary fiction that lead to stronger performance on ToM tasks?

Literary and popular fiction

Literary and popular fiction have long been differentiated in terms of the effort required by the reader to construct and infer meaning (e.g., Barthes, 1974; Bruner, 1986), and work in reader-response theory (Miall and Kuiken, 1994) and computational poetics (Jacobs, 2015) has highlighted structural differences in the amount of *foregrounding* (van Peer, 1986). Foregrounding refers to stylistic features at the phonetic, grammatical, and semantic level, that create various kinds of complexity and defamiliarize readers, forcing them to construct meaning—as opposed to simply receiving it. Compared to popular fiction, literary fiction is thought to disrupt the schematic knowledge that individuals use to make sense of interactions in daily life and readers utilize to make sense of fictional characters (Culpeper, 2001). One consequence of these discourse deviations (Cook, 1994) is that they render characters more complex and less predictable, forcing the reader into an interpretative and attributional effort (Culpeper, 1996) that is thought to be responsible for the observed effects of literary texts on ToM.

Both popular and literary fiction require interpretation, inferential, and attributional processes (Graesser et al., 1994), but *literary* characters "make[s] the reader infer implied mental states in addition to (and sometimes instead of) spelling some out" (Zunshine, 2019, p. 5). The effect of literary fiction on ToM may thus be due to the mentalizing effort required by its complex, unpredictable characters; characters often referred to as "round," in contrast to the "flat" characters of popular fiction (Forster, 2002). Recent empirical research confirms that characters of literary fiction are perceived as more complex than those of popular fiction (Kidd and Castano, 2019). Here I propose that the distinction between popular vs. literary fiction distinction has a parallel in the film: Hollywood vs. art film.

Hollywood and art film

The distinction between Hollywood and art film has been questioned and complexified from a variety of perspectives (see, for instance, Gaines, 1992). It retains, however, face validity in as much as it is routinely used in common discourse, and is considered a valuable framework in cognitive film theory

(e.g., Bordwell et al., 1985; Cutting, 2016; Cutting and Armstrong, 2018; Kuhn and Schmidt, 2014; Smith et al., 1999).

Similar to popular fiction, the Hollywood film is centered around the plot and its characters are types that help to move the story along smoothly. Plots vary considerably from one Hollywood genre to another (Altman, 1999; Schatz, 1981), but what is important here is that they indeed follow well-established, culturally shared, and rehearsed plots. These plots have corresponding schemas in the mind of the spectators and such schemas ease the task of forming representations of the characters' mental states (Levin et al., 2013). The characters of Hollywood films are equally well-established. Some are genre-specific types that exist primarily to guide our understanding of the fictional world and are of little relevance or impact in our daily life (Eder, Jannidis, Schneider, 2016). Examples are the villain in Bond movies (or Bond himself), the castaway, or the warrior. Other types typically found in Hollywood movies refer to social identities and roles that we routinely encounter in everyday life: accountants, ethnic minorities, men/women, religious groups. The simplified representation and stereotyping of these social groups (e.g., the nerdy, socially awkward accountant, the devout catholic, or the fanatic Arab) can be highly problematic since they contribute to perpetuating specific attitudes and prejudices towards these group members which then spill over into real life (Shaheen, 2003). However, they also allow spectators to easily follow the plot, drawing expected inferences and filling the gaps by basking in their social knowledge and Theory of Society: a special-purpose modular capacity to understand others in terms of culturally transmitted information about group membership, for example, which social groups exist in one's culture and which stereotypes adhere to these groups (Hirschfeld et al., 2007, p. 451; Hirschfeld, 2006). It is adherence to these theories of societies that allow us, while watching a generic Hollywood ending, to know that the heroine will say Yes! to the protagonist's declaration of love (Choi, 2005).

How does the art film differ from the Hollywood film? The art film category has historically comprised rather different types of film, such as the German expressionism of the 1920s and Italian neorealism of the 1940s and 1950s. Since they were originally, primarily produced outside of Hollywood and the United States, these films are also referred to as international art cinema. In the following decades, art film meant primarily films that heavily reflected the specific director's idiosyncratic technical competence and distinguishable personality (Sarris, 2007). To this date, in the French and Italian film discourse, art films are often referred to as film d'auteur/film d'autore. While the meaning of art film has certainly changed over time (Ndalianis, 2007), it is still understood primarily in contrast to the Hollywood film. Bordwell (1985) argues that contrary to Hollywood film, certain types of the film "undermine our conviction in our acquired schemata, open us up to improbable hypotheses, and cheat us of satisfying inferences." (p. 47). If in the art film the ready-made, stereotypical inferences allowed by Theory of Society are of little use (Wollen, 1972), spectators are, in their search for meaning (Stein and Trabasso, 1985), forced to rely on Theory of Mind processes to understand a character's psychology (Choi, 2005; Culpeper, 1996; 2001; Gerrig and Allbritton, 1990; Eder et al., 2016; Vaage, 2010).

Not only is the art film more likely to focus on the inner life of its characters, but it also makes it tougher to draw inferences about such mental states. "The art cinema uses 'realistic'–that is, psychologically complex characters. […] whereas the characters of the classical narrative have clear-cut traits and objectives, the characters of art cinema lack defined desires and goals." (Bordwell, 1979, pp. 57–58; see also Bruner, 1950; Tan, 2013; Zunshine, 2007).

The distinction between 'realistic' characters in art film and 'genre dependent' ones in Hollywood film, is also discussed by

Schweinitz (Schweinitz, 2011), who juxtaposes individual characters and types. The former are psychologically complex and multi-faceted, while the latter is schematically reduced. When these characters appear on the screen "they are already complete: defined, weighed, and minted." (Eco, 1986, quoted in Schweinitz, 2011, p. 45).

The distinction between Hollywood and art film is also invoked, directly or indirectly, by scholars who have specifically discussed the process through which spectators relate to character's psychology. Smith (1995) argues that the art film suppresses information that would unambiguously explain a character's expression and invites the viewer to a process of a simulation of a character's inner life to a much greater extent than the Hollywood film, which by presenting redundant information about the character, reduces viewers' degrees of freedom in deciding what a character wants, beliefs, feels and will do next. Vaage (2010) proposes that mainstream, Hollywood film elicits empathizing through embodiment, and it is about mind-feeling; the dedramatized, art film fosters empathy through the imagination and is about mind-reading.

The present experiment

Based on findings that literary fiction elicits greater ToM than popular fiction, and parallel theorizing in film theory, I hypothesize that watching an art film induces a shift to a cognitive framework that emphasizes ToM strategies, resulting in greater performance on tests of ToM after viewing an art film as compared to a Hollywood film.

I further hypothesize that this effect, at least in part, might be due to the fact that characters of art films are more complex and less predictable than those of Hollywood films. Finally, I hypothesize a sequential mediation, so that complexity and predictability, in a sequential manner, would mediate the effect of film type on ToM. The sequential aspect of the mediation is a matter of logic: Unpredictable scenarios do not necessarily have to be complex, but complex scenarios are less predictable— I shall return to this point in the discussion.

These hypotheses were explored through an experiment, in which participants were randomly assigned to watch a clip of one amid several art or Hollywood films. After viewing the clip, participants were asked to rate the characters before their performance on ToM tests was assessed. The goal of the experiment is *not* to test whether exposure to films in the context of the experiment has long-term effects on participants' ToM. As discussed above when reviewing empirical work on the effects of literature, this is best done via other methodologies. The focus of the present experiment is on establishing whether watching art (vs. Hollywood) films primes and trains ToM and what mediates this effect. As a long tradition in cognitive science has shown, experiments are best suited for establishing the causal relationship, and this is the use I make of this methodology in this context.

Methods

Participants. I recruited 326 American participants online via Amazon MTurk, a crowdsourcing marketplace for work that is used extensively in behavioral research and has been proven to be a reliable source of good quality data (Crump et al., 2013), and paid $6 for their participation. Of these, 17 did not complete the survey; 54 reported playback problems with the video, and 12 watched less than 19 min of the clip. Further, participants were excluded for performing at or below chance on the RMET ($n = 4$). Participants who rated negative intention + negative outcome scenarios on the MJS as more permissible than the neutral intention + neutral outcomes scenarios ($n = 3$) were

excluded, as these responses indicate careless or extremely unusual responding. Participants who spent a long time on the film page (>3.5 SD), indicating a long distraction from the study ($n = 4$), were also excluded. The final sample ($N = 232$) had 112 participants in the art condition (48%), and 120 in the Hollywood condition (52%). The sample consisted of 128 (55%) female participants and one participant indicated "other." Gender was equally distributed across conditions: $\chi^2 = 0.44$, $p = 0.50$, and Age ($M = 33.71$; $SD = 10.58$), $t(230) = 0.54$, $p = 0.46$, did not vary across conditions. Education was also measured. 10% reported having completed High School, 37% Some College, 44% College, and 9% a Graduate Degree. Education did not vary across conditions: $\chi^2 = 0.11$, $p = 0.98$. Ethnicity was measured by asking participants to choose among the following categories: Asian, Black, Latino, White, Native American, or Other. Given that 78.8% indicated White, the variable was recoded as white vs. non-white. Ethnicity did not vary across conditions: $\chi^2 = 0.04$, $p = 0.83$. Participants also indicated their Major: Business (16%), Humanities (21%), Natural Science (21%), Social Science (15%), and other (27%).

I undertook a priori power analysis to justify the sample sizes using G*Power. I assumed the effect size of Cohen's $d = 0.40$ (approximately $f = 0.20$), which is "typical for social psychology as a whole" (Gervais et al., 2015, p. 849). Using the partial eta squared, $f = 0.20$ with an alpha = 0.05 and power of 0.80, G*Power estimated a total sample size of 200. I thus believe that the final sample of $N = 232$ was adequate to meet the objectives of this study.

The research protocol was approved by the Human Research Protection Program (HRPP) of The New School University. Informed consent was obtained from all participants.

Procedure

Film selection. Research comparing the effects of two media types has typically used only one or two stimuli per category. For instance, comparing one violent video game to a non-violent one. This is problematic if the goal is to draw conclusions beyond the specific stimuli, at the media category level. Given the goal of the experiment, I followed recent methodological guidelines with regard to the number of stimuli and their selection (Reeves et al., 2016). An initial list of 18 Hollywood films was randomly selected from a list of the top-25 highest-grossing films (worldwide) as reported by the *Yearly Box Office* database at BoxOfficeMojo.com; 15 Art films were selected from a list of winners and nominees of the Palme d'Or at the Cannes Film Festival as reported by the *Event History* database on the Internet Film Database. All films were released between the years 2000–2010, and to ensure a meaningful delineation between art and Hollywood films, monetary limits were set. Art films had to have grossed <$50 million and Hollywood films >$250 million (worldwide). The purpose of this cutoff was to exclude Hollywood "flops" that would not be the best representations of the genre, and to minimize the inclusion of the kinds of "fuzzy" art films that are relatively more successful than their peers (other Cannes nominees) that potentially indicate some cross-over in stylizations (e.g., Quentin Tarantino's *Inglorious Bastards*, Sophia Coppola's *Marie Antoinette*, or Roman Polanski's *The Pianist*). Through evaluation and discussion with students enrolled in the course *Introduction to Screen Studies* at a liberal arts college, and with experts, six art and six Hollywood films were selected and used for this study. Art films: Spider (2002), Clean (2006), Certified Copy (2011), Bright Star (2009), All or Nothing (2002), Synecdoche, New York (2008); Hollywood Films: Monster in Law (2005), Fast and Furious (2009), Click (2006), Pirates of the Caribbean (2006), Quantum of Solace (2008), The Dark Knight (2008). The first 20 min of each of these

films were used as stimuli. No mention of the distinction between art and Hollywood films was made, and participants were not told whether the clip they were assigned to watch was of a film categorized as art or Hollywood.

Qualtrics was used for stimuli presentation and the recording of all data. After giving consent, assuring confidentiality and anonymity, participants were informed that to continue with the study they should be able to watch a clip of approximately 20 min and to use headphones while viewing it in full-screen mode. They were randomly assigned to watch one of 12 clips (6 art, 6 Hollywood). The clips were hosted on the video platform Vzaar, an online video hosting service (because their file sizes were too large for Qualtrics' video players limits) and accessed automatically by Qualtrics to be shown to participants. Participants then completed a series of tasks.

Film evaluation. Hollywood films are expected to be evaluated more positively to be easier to understand than art films, and also more likely to have been previously seen by participants. Two evaluation items were measuring enjoyment of the film and desire to watch it its entirety correlated strongly [$r(232) = 0.88$] and were thus averaged into a composite score ($M = 5.56$, $SD = 1.73$). Participants also indicated how easy to understand the film was and whether they had seen it before. Aside from the last question, which was answered as yes or no, the three other items were answered on a scale from 1 to 7, with different endpoint labels depending on the question.

Character assessment

Viewers focus their attention on the film's main characters (Magliano et al., 2005). Upon viewing the film excerpt, participants were thus asked to judge the main character. Participants rated how *complex, a type, obscure,* and *behaving the same across situations,* the main character was. They also judged the main character on five semantic differential items (ugly/beautiful, bad/good, unpleasant/pleasant, dishonest/honest, awful/nice) and 10 items from the Big Five inventory (e.g., tends to be lazy, is outgoing, sociable). These questions provided the context to measure character *predictability* by asking participants how confident they were in their assessments and the extent to which they could predict the character's future behavior. All questions were answered on a scale from 1 to 7, with different endpoint labels depending on the question.

Theory of mind

The first measure of ToM used was the *Moral Judgment Task* (Young et al., 2007). This task consists of scenarios in which an actor with neutral or negative intentions brings about an outcome for another character that can be neutral or negative in nature (for example, see Fig. 1 in Young et al., 2007). The abbreviated version used here (Young et al., 2010) comprises 3 scenarios for each of the combinations of intentions and outcomes, for a total of 12 scenarios. For each scenario participants answered on a scale from 1 (forbidden) to 5 (permissible) the action performed in the scenario. From participants' answers, two scores are computed. *Moral Mind* scores are calculated by subtracting moral permissibility ratings for failed harms from those of accidental harms. This score is higher the more participants take into consideration the intentions of the actors, which is a measure of ToM. *Moral Base* scores are calculated by subtracting the moral permissibility ratings of intended harms from those for neutral scenarios (Young et al., 2007). A score can thus be computed by subtracting Moral Base from Moral Mind. Higher scores are to be interpreted as the stronger capacity to consider the intentions of the actor, and thus stronger ToM performance (Young et al., 2007). Scores tend to be negative because judgments of the permissibility of the Moral Base are higher than that of Moral Mind, and so it was in the present sample, MJS: $M = -1.72$, $SD = 1.49$.

The second measure was the Reading the Mind in the Eyes Test (RMET). This task comprises 36 trials in which an image of an actor's eyes is shown, and the participant must choose which of four complex emotion terms (e.g., sympathetic, irritated, thoughtful, encouraging) best matches the actor's mental state. The RMET is scored by summing the correct responses, whereby "correct" refers to responses given by judges and a large sample of adults in validation studies (Baron-Cohen et al., 2001). The RMET has been used as a measure of ToM in experiments that varied the type of fiction participants read (e.g., Black and Barnes, 2015a; Kidd and Castano, 2013) or the type of TV drama they watched (Black and Barnes, 2015b). The RMET Mean (26.15) and SD (4.50) were comparable with those from these previous studies.

Participants were also asked whether they experienced any playback problems with the video (e.g., skipping, sound problems, etc.), and to indicate gender and age. Participation took ~45 min.

Results

Film evaluation. Following recommendations that variance due to individual stimuli representing the different categories of interest be factored in and treated as random factors (e.g., Reeves et al., 2016; J. Brunner, personal communication, February 1, 2019), I used a mixed model with Type of Film (art vs. Hollywood) as fixed factor and Film as a random factor to analyze the film evaluation score. In line with the difference at the box office, compared to Hollywood films, art films were evaluated less positively. The same analysis revealed that art films were judged to be less easy to understand. Results are presented in Table 1. Also, participants were more likely to report having seen the movie in the popular film (33%) than in the art film (3.57%) condition.

Theory of mind (ToM). The exact same model described above was used to test the effect of the type of film on the two ToM measures, the Moral Judgment Score and the Reading the Mind in the Eyes Test. In both instances, performance was higher in the art film, compared to the Hollywood film condition. See Table 1. The two measures showed a significant, positive correlation, $r(232) = 0.24$, $p < 0.001$. The means and SDs of the two ToM measures for each film are presented in Table S1.

Character assessment. A factor analysis with oblique rotation was conducted on the Predictability and Complexity items. This revealed the presence of two factors. On the first factor (labeled Predictability), the two items asking about confidence in the knowledge of the character and about predicting the character's behavior in the future loaded strongly (>0.80) and uniquely. On the second factor, (labeled Complexity) three items (type, complex, obscure) loaded strongly (>0.74) and uniquely. One item (the character behaves the same across situations) had loadings <0.40 on both factors and was thus excluded from further analysis. I then averaged the respective items to form Complexity ($M = 4.20$, $SD = 1.31$) and Predictability ($M = 5.36$, $SD = 1.05$) scores. These scores were entered as dependent variables in the same mixed model described above. While means went in the expected directions for both Complexity and Predictability, the results were significant only for Predictability—see Table 1. As expected, Complexity and Predictability correlated negatively, $r(232) = -0.17$, $p = 0.008$.

Mediation. I tested the sequential mediational model (Type of film > Complexity > Predictability > ToM) using the PROCESS bootstrapping method to test indirect effects (i.e., mediation) with a confidence level set at 0.95 and bootstrap bias-corrected samples set at 10,000 (Model 6, Preacher and Hayes, 2004). Results are shown in Fig. 1a. For ToM measure, the strength of the indirect (MJS: $b = 0.004$, SE = 0.003, 95% CI [0.001, 0.015]; RMET: $b = 0.006$, SE = 0.006, 95% CI [0.001, 0.018]) and of the direct effect (MJS: $b = 0.1255$, SE = 0.0674, 95% CI [−0.007, 0.258];

RMET: $b = 0.063$, SE = 0.067, 95% CI [−0.069, 0.196]) suggested that the mediating factors accounted for much of the effect of Type of film. The same sequential model in which the position of Complexity and Predictability was reversed (Type of film > Predictability > Complexity > ToM) did not indicate mediation (as indicated by the indirect effects' CIs containing the value zero), for both RMET and MJS.

An alternative to the proposed model (Fig. 1b) consists of having Complexity and Predictability as parallel mediators (that is, a model in which Type of Film impacts both variables and independently mediates the effect of Type of Film on ToM). I thus tested this alternative model for both MJS and RMET. For MJS, complexity did not emerge as a reliable mediator ($b = -0.006$, SE = 0.014, 95% CI [−0.038, 0.019]) while predictability did ($b = 0.033$, SE = 0.017, 95% CI [0.006, 0.076]). For RMET complexity did not emerge as a reliable mediator ($b = -0.012$, SE = 0.014, 95% CI [−0.047, 0.011]) while predictability did ($b = 0.042$, SE = 0.020, 95% CI [0.0011, 0.095]).

Supplementary analysis. The feedback I received on earlier drafts of this article led us to consider further analyses. Since they were not planned in advance, I report them in this section, as supplementary.

Hollywood films were evaluated more positively than art films. This could be considered a manipulation check since Hollywood films are made with the primary goal to entertain, to be enjoyable, and thus to be popular (Harris et al., 2017). Nonetheless, it is informative to explore whether evaluation accounts for any of the effects on MJS, RMET, Complexity and Predictability. I thus tested the exact same models described above to assess the effects of Type of film on these dependent variables and mediators, adding film evaluation as a covariate. The only result that changed was for Predictability, with the effect of Type of film slightly reduced ($F = 3.69$, $p = 05$ to $F = 2.47$, $p = 0.15$). I thus also re-run the mediational analyses, using the exact same models as described above, but adding film evaluation as a covariate. The results were unchanged.

Hollywood films were also judged as easier to understand than art films. Using the same covariate approach described above, I thus checked whether this variable accounts for any of the effects on MJS, RMET, Complexity and Predictability. Adding it as a covariate caused a change only on the Predictability finding, ($F = 3.69$, $p = 05$ to $F = 0.38$, $p = 0.53$) and rendered the indirect effect of Type of film in the mediational models for both the RMET and MJS, no longer significant.

I also looked at IMDB ratings—publicly available 1–10 film evaluations provided by registered users. This is of course a film-level variable, not an individual-level one. Only a marginal difference between art and Hollywood films emerged on IMDB ratings, $F(1, 230) = 3.14$, $p = 0.08$. While this may at first appear as inconsistent with the evaluation results described above, it should be noted that the IMDB ratings are by individuals who

	Means		E	SE	CI		F	p
	Art film	**Hollywood film**						
MJS	−1.50	−1.94	0.45	0.20	0.05	0.85	4.88	0.03
RMET	26.58	25.73	0.84	0.42	0.01	1.67	4.02	0.04
(Character) Complexity	4.46	3.94	0.52	0.40	−0.28	1.31	1.63	0.20
(Character) Predictability	5.12	5.56	−0.44	0.23	−0.89	0.01	3.69	0.05
(Film) Evaluation	5.03	6.01	0.97	0.31	−1.06	−0.35	9.55	0.01
(Film) Understandable	5.39	6.45	−1.06	0.30	−1.67	−0.45	11.9	0.01

Table 1 Effects of type of film (Art vs. Hollywood film).

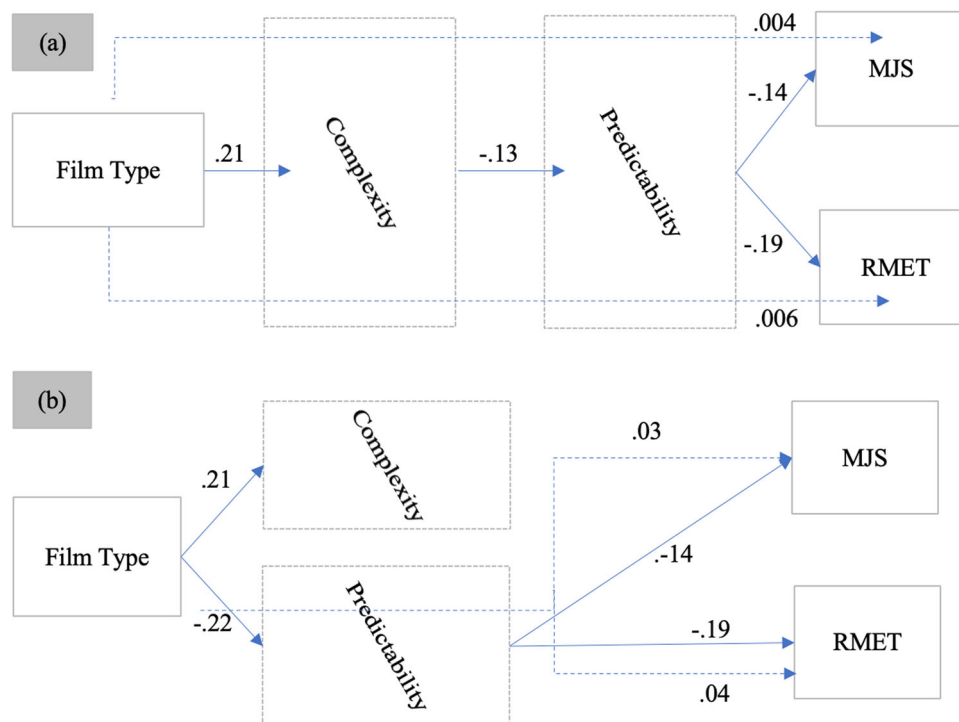*MJS moral judgment score, RMET reading the mind in the eyes test.*

Fig. 1 Film Type: art = 1, Hollywood = −1; MJS = moral judgment score; RMET = reading the mind in the eyes test. Coefficients reported are unstandardized **b**. Confidence intervals of all effects reported, do not include zero. Dotted arrow lines refer to indirect effects. Proposed sequential model (**a**) and alternative parallel model (**b**).

chose to watch the film, while the film evaluation in the present dataset is provided by participants who were randomly assigned to watch one of the films. IMDB ratings did not correlate with the RMET ($r = 0.01$, $p = 0.85$), the MJS ($r = 0.01$, $p = 0.94$), or character Predictability ($r = 0.02$, $p = 0.71$). They correlated, however, with character Complexity, ($r = 0.30$, $p = 0.01$). The higher the IMDB score, the greater the perceived complexity of the main character. Controlling for IMDB ratings in the analyses reported above, which included Complexity, did not change the results. The results of the mediational model for both the RMET and MJS were also unchanged when controlling for IMDB.

As reported above, only 3.5% of participants in the art film condition reported having seen the movie before, while 33% of those in the Hollywood film did. I dummy coded this variable (Seen, yes = 1; no = −1) and computed correlations. Having seen the movie or not did not correlate with the RMET ($r = −0.03$, $p = 0.58$), the MJS ($r = 0.01$, $p = 0.90$), or character Complexity ($r = −0.05$, $p = 0.43$). It showed a small but significant correlation with character Predictability, ($r = 0.16$, $p = 0.01$). As I did in the cases described above where a correlation emerged with a variable of interests, I computed the same analyses adding Seen as a covariate. The previously reported effect of the Type of film on predictability changed slightly ($F = 3.69$, $p = 0.05$ to $F = 2.90$, $p = 0.08$), but results of the mediational model for both the RMET and MJS were unchanged.

I also tested whether gender impacted the ToM measures, but found no reliable effect on either the RMET ($F = 0.71$, $p = 0.39$) or the MJS ($F = 2.16$, $p < 0.14$). Adding gender as a covariate in the mediational models did not alter the findings.

College major had a significant impact on the RMET, $F(4, 227) = 2.53$, $p = 0.04$, due primarily to the higher scores among Humanities ($M = 27.53$) majors, compared to Business ($M = 26.05$), Natural Science ($M = 26.40$), Social Science ($M = 26.22$) and Other ($M = 24.85$). A contrast pitting Humanities

vs. all other majors was significant, $F(1, 227) = 5.24$, $p = 0.02$. This pattern closely replicates that obtained in previous research (Kidd and Castano, 2017a). A chi-square analysis confirmed that random assignment had resulted in equal distributions of the various majors across conditions, X^2 (4, $N = 232$) = 1.56, $p = 0.81$. Adding major as a covariate to the analyses reported above for RMET did not change the results. The effect of major was not significant for MJS.

Finally, education had no effect on the MJS, but a marginal effect on RMET, $F(3, 228) = 2.35$, $p = 0.07$, driven primarily by the Graduate Degree group ($M = 28.55$), compared to the College ($M = 25.65$), Some College ($M = 26.12$) and High School ($M = 26.33$) groups. However, given that the Graduate Group was only 9% of the sample, this result should be taken with a grain of salt. Adding education as a covariate to the analyses reported above for RMET did not change the results.

Discussion
In this article, I propose that viewing art films, compared to Hollywood films, results in greater performance on Theory of Mind tasks, and that this effect is mediated by how film characters are perceived by the viewer. The results of an experiment in which participants were randomly assigned to watch excerpts of one of six art versus one of six Hollywood films, support this proposition. Compared to those who viewed a Hollywood film, those who viewed an art film scored higher on two correlated, yet distinct measures of ToM: the RMET, which is a socio-perceptual measure that relies on the processing of visual information, and the MJS, which is a socio-cognitive measure that relies on the representation of actors in written fictional scenarios.

The results of the experiment further showed that art-film characters are perceived as less predictable than those of Hollywood films—a finding which is in line with the much-theorized ambiguity and uncertainty of art films (Bordwell and Thompson,

1993). The difference in the perceived complexity of characters, if not below the conventional threshold of statistical significance, was in the expected direction. Finally, mediational analyses provided support for the sequential process: Type of film > Complexity > Predictability > ToM. This model is the most logical one since it stands to reason that complex things are less easily predictable than simple ones. I also tested models with alternative mediational patterns. The one reversing the sequence of the mediators was not supported, while the one testing the parallel mediation received only partial support.

The pattern of results was largely independent of other variables at the participant-level or film-level of analysis. The extent to which participants found the film to be easy to understand also did not alter the effects of film type on the ToM measures, but when this was added as a covariate in the mediational models, it made the indirect effect no longer significant. The fact that art films were judged as less easy to understand (and that participants who found the film they watched difficult to understand also tended to see the character of the movie as less predictable) further strengthens the rationale concerning the complexity of art films and their characters.

In support of the film taxonomy, Hollywood films were liked more by participants in the experiment (individual-level) and had higher IMDB ratings (film-level), than art films, but neither of these variables accounted for the results on ToM. Similarly, having seen the film did not matter much when it comes to the effects of film exposure within the experiment on ToM. This is an informative finding for future research aiming at comparing the effects of classes of cultural products that may have different diffusion in the population.

Finally, I replicated previous findings documenting the effect of college majors on RMET performance (Kidd and Castano, 2019), with Humanities majors scoring the highest. In the present experiment, different majors were equally distributed across the two conditions, and adding major as a covariate to the main analyses did not modify the results.

The pattern of findings shows that individuals that watched art films performed better on two ToM tasks, compared to those who watched Hollywood films. As noted above, results of experiments such as this aim at establishing possible causal relationships and should not be interpreted as evidence that short exposure to a type of film (or novel) causes stable changes in socio-cognitive skills for which longer and repeated exposure is necessary. Research on cognitive processes in young children has shown the parallel between the effects obtained immediately after a single administration of a quick brain-train exercise and the longer-term effects that follow from repeated and longer brain-train exercises over four months (Wexler et al., 2016; see also Batini et al., 2021). In adults, due to their lessened plasticity (Baltes and Kliegl, 1992), it might be more difficult to produce reliable and stable changes in a four-month period, and longer/more intense research interventions might be difficult to carry out. Recent findings, however, suggest that brain connectivity in areas associated with ToM is affected by fiction reading and that some changes are still observable several days after the reading has occurred (Berns et al., 2013; see also Bartolucci and Batini, 2019). Alternative strategies to assess the long-term impact of watching art vs. Hollywood film are cross-sectional correlational studies similar to those that have shown the impact of lifetime exposure to literary fiction on ToM (e.g., Kidd and Castano, 2017a), and longitudinal studies.

An additional question, which is related to the above discussion, concerns the nature of the effects on ToM that emerged in the present experiment, as well as in previous experiments focusing on viewing TV drama (Black and Barnes, 2015b) playing videogames (Bormann and Greitemeyer, 2015) or reading books

(Kidd and Castano, 2013). What do we mean when we say that some of these experiences foster ToM? Do participants get better at ToM or the enhanced performance on ToM tasks is a byproduct of other factors such as the propensity to engage in mentalizing or other cognitive skills that may scaffold ToM skills (Apperly, 2012). The present experiment was not designed to provide an answer to the latter question, which is ultimately part of a broader question of domain-specific versus domain-general processes (Heyes, 1998; Saxe et al., 2006; Tomasello, 2010). What I propose happens while watching an art film is a shift to a mode of social cognition that privileges ToM processes as opposed to relying primarily on Theory of Society, which is considered the default mode of social perception (Brewer, 1988; Fiske and Neuberg, 1990). This shift is tantamount to a prime, which results in greater performance on ToM tasks carried out immediately after. As discussed above, repeated exposure may result in stable changes in both the propensity to use ToM processes and in the accuracy achieved by these processes.

To test the effects of viewing art vs. Hollywood films I used 12 different films as stimuli and used individual films as a random factor in the analyses. A recent review of studies in media psychology (Reeves et al., 2016) indicates that amid studies on the effects of types of media indexed in google scholar, 60% used only one stimulus per condition, and only 13% of studies used four or more stimuli. Amid studies using films as stimuli, no study used more than two stimuli. Just as important, no study analyzed stimulus repetition as a random factor (the recommended practice; Reeves et al., 2016). Particularly in light of this data, the methodology and analytical strategy used in the present experiment constitutes a significant improvement upon much research in media psychology, including recent work that has focused on the impact of TV drama (Black and Barnes, 2015b) or videogames (Bormann and Greitemeyer, 2015) on Theory of Mind. In spite of such improvements, the experiment presented here is only a first step toward testing the hypothesis that art vs. Hollywood films differently impacts social cognition processes. Future research may provide sharper, more refined categorizations of cinematic experiences, that better account for the differences I found.

During the review process, three points have been raised with regard to the analytical strategy I adopted, which are worth mentioning. One points to the fact that, compared to the art film condition, a higher proportion of participants in the Hollywood film condition reported having already seen the film. It could be argued that previous familiarity with the characters, instead of the quality of the characters themselves, might be partially responsible for the effect. I believe that statistically controlling for this variable is the appropriate analytical strategy. However, future research could substantially oversample in the Hollywood film condition, and exclude from the analyses all participants who reported having seen the film. A second point concerns the exclusion of some participants based on their response pattern. In the extensive literature on the RMET, criteria for exclusion have varied, and with regard to the moral judgment task, the literature is much thinner. In this study, in which these two ToM measures were used, the same exclusion criteria used in the pre-registered studies by Kidd and Castano (2019), which were themselves based on previous research (e.g., Chapman et al., 2006), and in most recent research on the RMET (e.g., Eddy and Hansen, 2020), were utilized. This seems an appropriate decision given that this study was conceptually and empirically modeled after Kidd and Castano's (2019) research studies. A third point concerns the rationale for the direction of causality between the perceived complexity and predictability of the characters—and the mediational model that followed. It could be argued that characters that are unpredictable are in turn perceived as

complex. This is not unreasonable, but I stand by the rationale presented above, namely that scenarios do not need to be complex to be unpredictable, but complex scenarios make for less certain predictions. In discussions of these findings that I have had with various colleagues, I was often asked about characters of Hollywood movies that transform themselves, in a rather unexpected, surprising fashion, over the course of the film. Although I am certainly not arguing that Hollywood characters cannot be unpredictable and complex, I would argue that such transformations typically entail moving from *type* to *type*: from honest to dishonest, from cynical to empathic, from good too bad—or vice versa, of course. These transformations are at least in some cases unpredictable, but it does not mean that the characters are complex. They are either, or.

Notwithstanding the fact that I believe the adopted analytical strategy is appropriate, I recognize that it is debatable, and that future research is needed to assess the robustness of the findings and to explore the many other avenues that the findings suggest.

One such avenue is the structural characteristics of films (or realization, Visch and Tan, 2008). Average shot length appears to be an interesting characteristic to consider. It has already attracted the attention of cognitive psychologists interested in film (Cutting, 2016), and maybe related to the speed of movement, which in turn is related to the attribution of mind to the moving targets (Morewedge et al., 2007). A second characteristic is misordering, namely ordering scenes in a manner that does not correspond to the natural course of events (Levin et al., 2013). Another characteristic is the type of shot. Shot scale can moderate viewers' use of ToM processes with regard to film characters (Plantinga, 1999; Rooney and Bálint, 2018). I argue, however, that facilitating viewers' access to characters' facial expression facilitates attunement and entrainment processes (Kinsbourne, 2005), possibly through mirror networks (Keysers, 2011), that are automatic and effortless, and thus do not require ToM processing (Martingano, 2020; Martingano and Castano, 2020). A similar point can be made for point-of-view shots (POVs; for a description, see Bordwell and Thompson, 1993).

POVs have been theorized to focus the attention of the spectator on the mental state of the character (Cutting and Armstrong, 2018; Levin et al., 2013). Film scholars, however, suggest that POVs reduce the need for mentalizing because they give direct access to the subjective experience of the characters (Choi, 2005). Instead of eliciting mindreading and imaginative empathy, they foster embodied empathy because they facilitate "latching onto the character's state through automatic mental mechanisms such as mimicry and feedback." (Vaage, 2010; p. 163). In these cinematic experiences, the epistemic gap between the viewer and the character is diminished, and much less imagining and inferring is required to the viewer.

Anecdotal evidence suggests that POVs are more frequently found in Hollywood than in art films. In her analysis of art-house film *Jeanne Dielman*, Choi (2005) discusses how Akerman, the film director, makes no use of point-of-view or reaction shots, and instead "presents the character's psychology only indirectly, that is, via visual disturbances and the pace of movements, rather than through the cause and effect linear logic often found in Hollywood narratives." (p. 24). Similarly, Vaage (2010) juxtaposes mainstream films, in which subjective narration elicits the feeling of being the character, to dedramatized film, which makes little use of POVs and close-ups and requires a more imaginative engagement by the viewer. I concur with Choi's (2005) and Vaage's (2010) somewhat counterintuitive rationale: it is cinematic experiences that withhold information about a characters' psychology that force viewers to engage in imagining and inferring processes. Films that impede entrainment and attunement

with, and imitation of, characters, are those that may paradoxically foster ToM skills. In fact, experimental research shows that when imitation is inhibited, ToM is enhanced (Santiesteban et al., 2012). Future research will help clarify whether POVs and other structural characteristics of films play a role in the pattern of results presented here.

The findings presented here complement other lines of research comparing how individuals process narrative in written fiction vs. film. While important differences exist between written fiction and films (Kuhn and Schmidt, 2014), research has shown that indexing strategies adopted by viewers to index films (Magliano et al., 2001) are similar to those used for narrative text (Zwaan et al., 1995) and that spectators/readers track the goals of characters in similar ways for the film (Magliano et al., 2005) and written narrative (Suh and Trabasso, 1993). Recent neuro-imaging research also found a similar pattern of brain activation, including ToM areas, when participants watched a movie or read its screenplay text (Tikka et al., 2018). The results also suggest that narrative can impact social cognition processes regardless of the medium through which it is delivered, but they go further and highlight a parallel between (written) literary fiction and art film, and (written) popular fiction and Hollywood film.

Most research conducted so far on the impact of fiction on social cognition has focused on the impact of written fiction on the Theory of Mind. Recent research has expanded this focus and showed that lifetime exposure to literary fiction positively predicts attributional complexity, while exposure to popular fiction negatively predicts it, and that exposure to literary but not popular fiction predicts greater social accuracy and lower egocentric bias (Castano et al., 2020). I would expect that Hollywood and art films differently impact these social cognition processes and cognitive styles in a way that is similar to popular and literary fiction, respectively.

I found that art (vs. Hollywood) films elicited stronger performance to ToM tests. This is not to be interpreted as art films being superior to Hollywood films. Storytelling served a variety of functions in our evolutionary history and continues to do so. One of these functions is fostering the Theory of Mind (Wiessner, 2014). Another, possibly primary function, is to increase group cohesiveness, sense of belonging, and cooperation (Smith et al., 2017), through creating, maintaining, and spreading the Theory of Society. This latter function is probably better achieved by the type of stories that are considered popular (Kidd and Castano, 2013)—and by the Hollywood film. To the extent that it does not reproduce negative stereotypes and foster prejudicial attitudes, the Theory of Society is just as important, from a societal standpoint, as Theory of Mind. At least in this regard, therefore, to propose a hierarchy of films, just as a hierarchy of fiction, seems a futile endeavor.

Data availability
The dataset is available from the corresponding author on reasonable request.

References
Altman R (1999) Film/genre. British Film Institute, London
Eder J, Jannidis F, Schneider R (eds) (2016) Characters in fiction: an introduction. revisionen: characters in fictional worlds: under- standing imaginary beings in literature, film, and other media. Walter de Gruyter, Berlin, Germany
Apperly IA (2012) What is "theory of mind"? Concepts, cognitive processes and individual differences. Quart J Exp Psychol 65(5):825–839

Baltes PB, Kliegl R (1992) Further testing of limits of cognitive plasticity: negative age differences in a mnemonic skill are robust. Dev Psychol 28(1):121–125. https://doi.org/10.1037/0012-1649.28.1.121

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. J Child Psychol Psychiatry 42(2):241–251. https://doi.org/10.1111/1469-7610.00715

Barthes R (1974) S/Z (R. Miller, Trans.) Hill and Wang

Bartolucci M, Batini F (2019) Long term narrative training can enhance cognitive performances in patients living with cognitive decline. Educ Gerontol 45:469–475. https://doi.org/10.1080/03601277.2019.1658384

Batini F, Lpuerini V, Cei E, Izzo D, Toti G (2021) The association between reading and emotional development: a systematic review. J Educ Training Stud 9(1):12–50

Berns GS, Blaine K, Prietula MJ, Pye BE (2013) Short-and long-term effects of a novel on connectivity in the brain. Brain Connect 3(6):590–600

Black JE, Barnes JL (2015b) Fiction and social cognition: the effect of viewing award-winning television dramas on theory of mind. Psychol Aesthet Creat Arts 9(4):423

Black JE, Barnes JL (2015a) The effects of reading material on social and non-social cognition. Poetics 52:32–43

Bordwell D (1979) The art cinema as a mode of film practice. Film Criticism 4 (1):56–64

Bordwell D (1985) Narration in the Fiction Film. University of Wisconsin Press

Bordwell D, Staiger J, Thompson K (1985) The classical Hollywood cinema: film style and mode of production to 1960. University Press, Columbia

Bordwell D, Thompson K (1993) Film art: an introduction. McGraw-Hill

Bormann D, Greitemeyer T (2015) Immersed in virtual worlds and minds: effects of in-game storytelling on immersion, need satisfaction, and affective theory of mind. Soc Psychol Personal Sci 6(6):646–652

Brewer MB (1988) A dual process model of impression formation. In:Wyer Jr RS, Srull TK (eds) Adv Soc Cogn. Erlbaum, Hillsdale, pp. 1–36

Bruner J (1950) Social psychology and group processes. Ann Rev Psychol 1 (1):119–150. https://doi.org/10.1146/annurev.ps.01.020150.001003

Bruner J (1986) Actual minds, possible worlds. Harvard University Press

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

Chapman E, Baron-Cohen S, Auyeung B, Knickmeyer R, Taylor K, Hackett G (2006) Fetal testosterone and empathy: evidence from the empathy quotient (EQ) and the "reading the mind in the eyes" test. Soc Neurosci 1(2):135–148

Choi J (2005) Leaving it up to the imagination: POV shots and imagining from the inside. J Aesthet Art Criticism 63(1):17–25

Cook G (1994) Discourse and literature: the interplay of form and mind. Oxford UP, Oxford

Crump MJC, McDonnell JV, Gureckis TM (2013) Evaluating Amazon's mechanical turk as a tool for experimental behavioral research. PLoS ONE 8(3): e57410. pmid:23516406

Culpeper J (1996) Inferring character from texts: attribution theory and foregrounding theory. Poetics 23(5):335–361

Culpeper J (2001) Language and characterisation: people in plays and other texts. Longman, London

Cutting JE (2016) The evolution of pace in popular movies. Cogn Res 1(1):1–21. https://doi.org/10.1186/s41235-016-0029-0

Cutting JE, Armstrong KL (2018) Cryptic emotions and the emergence of a metatheory of mind in popular filmmaking. Cogn Sc 42(4 May):1317–1344. https://doi.org/10.1111/cogs.12586

Decety J, Jackson PL (2016) A social-neuroscience perspective on empathy. Curr Dir Psychol Sci 15(2):54–58. https://doi.org/10.1111/j.0963-7214.2006.00406.x

Eco U (1986) Die praktische Anwendung der literarischen. In: Person UE (ed) Apokalyptiker und Integrierte. Zur kritischen Kritik der Massenkultur, Frankfurt/M, pp. 161–186

Eddy CM, Hansen PC (2020) Predictors of performance on the Reading the Mind in the Eyes Test. PLoS ONE 15(7):e0235529

Fiske ST, Neuberg SL (1990) A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. Adv Exp Soc Psychology 23:1–74

Forster EM (2002) Aspects of the Novel. RosettaBooks (1927)

Frith U (1989) Autism: Explaining the Enigma. Basil Blackwell

Gaines J (1992) Classical Hollywood Narrative: The Paradigm Wars. Duke University Press

Gallese V, Goldman A (1998) Mirror neurons and the simulation theory of mind-reading. Trends Cogn Sci 2(12 Dec):493–501. https://doi.org/10.1016/s1364-6613(98)01262-5

Gerrig RJ, Allbritton DW (1990) The construction of literary character: a view from cognitive psychology. Style pp. 380–391

Gervais WM, Jewell JA, Najle MB, Ng BK (2015) A powerful nudge? Presenting calculable consequences of underpowered research shifts incentives toward adequately powered designs. Soc Psychol Personal Sci 6(7):847–854

Goldstein TR, Winner E (2012) Enhancing empathy and theory of mind. J Cogn Dev 13(1):19–37

Gopnik A, Wellman HM (1992) Why the child's theory of mind really is a theory. Mind Lang 7(1-2):145–171. https://doi.org/10.1111/j.1468-0017.1992.tb00202.x

Gordon R (1986) Folk psychology as simulation. Mind Lang 1(2):158–171

Graesser AC, Singer M, Trabasso T (1994) Constructing inferences during narrative text comprehension. Psychol Rev 101(3 Jul):371–395. https://doi.org/10.1037/0033-295x.101.3.371

Harris RJ, Borror SW, Koblitz KR, Pearn M, Rohrer TC (2017) Memory for emotional content from entertainment film: the role of subtitles, sex, and empathy. Media Psychol 20(1):28–59

Heyes CM (1998) Theory of mind in nonhuman primates. Behav Brain Sci 21 (1):101–114

Heyes CM, Frith CD (2014) The cultural evolution of mind reading. Science 344 (6190 Jun):1243091. https://doi.org/10.1126/science.1243091

Hirschfeld L (2006) Who needs a theory of mind. Biological and cultural bases of human inference 131:160

Hirschfeld L, Bartmess E, White S, Frith U (2007) Can autistic children predict behavior by social stereotypes? Curr Biol17(12):R451–R452

Jacobs AM (2015) Neurocognitive poetics: methods and models for investigating the neuronal and cognitive-affective bases of literature reception. Front Human Neurosci 9:186. https://doi.org/10.3389/fnhum.2015.00186

Kanske P, Böckler A, Trautwein FM, Singer T (2015) Dissecting the social brain: Introducing the EmpaToM to reveal distinct neural networks and brain–behavior relations for empathy and Theory of Mind. NeuroImage 122:6–19

Keysers C (2011) The empathic brain: How the discovery of mirror neurons changes our understanding of human nature. Lulu.com

Kidd DC, Castano E (2013) Reading Literary Fiction Improves Theory of Mind. Science 342(6156 Oct):377–380. https://doi.org/10.1126/science.1239918

Kidd DC, Castano E (2017a) Different stories: how levels of familiarity with literary and genre fiction relate to mentalizing. Psychol Aesthet Creat Arts 11 (4):474–486. https://doi.org/10.1037/aca0000069

Kidd DC, Castano E (2017b) Panero et al. (2016): Failure to replicate methods caused the failure to replicate results. J Personal Soc Psychol 112(3):e1–e4. https://doi.org/10.1037/pspa0000072

Kidd DC, Castano E (2019) Reading literary fiction and theory of mind: three pre-registered replications and extensions of Kidd and Castano (2013). Soc Psychol Personal Sci 10(4):522–531. https://doi.org/10.1177/1948550618775410

Kidd DC, Ongis M, Castano E (2016) On literary fiction and its effects on theory of mind. Scient Study Literat 6(1):42–58. https://doi.org/10.1075/ssol.6.1.04kid

Kinsbourne M (2005) Imitation as entrainment: brain mechanisms and social consequences. Perspect Imitation 2:163–172

Kuhn M, Schmidt JN (2014) Narration in film. Handbook of Narratology, Berlin, pp. 384–405

van Kuijk I, Verkoeijen P, Dijkstra K, Zwaan RA (2018) The effect of reading a short passage of literary fiction on theory of mind: a replication of Kidd and Castano (2013). Collabra 4(1 Feb):1–12. https://doi.org/10.1525/collabra.117

Levin DT, Hymel AM, Baker L (2013) Belief, desire, action, and other stuff: theory of mind in movies. In A. Shimamura (Ed.), Psychoncinematics (pp. 244–266). New York, NY: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199862139.003.0013

Magliano JP, Miller J, Zwaan RA (2001) Indexing space and time in film understanding. Appl Cogn Psychol 15(5):533–545

Magliano JP, Taylor HA, Kim HJJ (2005) When goals collide: monitoring the goals of multiple characters. Memory Cogn 33(8):1357–1367

Mar RA, Oatley K (2008) The function of fiction is the abstraction and simulation of social experience. Perspect Psychol Sci 3(3):173–192

Mar RA, Oatley K, Peterson JB (2009) Exploring the link between reading fiction and empathy: Ruling out individual differences and examining outcomes. Communications 34(4):407–428. https://doi.org/10.1515/COMM.2009.025

Mar RA, Oatley K, Hirsh J, dela Paz J, Peterson JB (2006) Bookworms versus nerds: exposure to fiction versus non-fiction, divergent associations with social ability, and the simulation of fictional social worlds. J Res Personal 40 (5):694–712. https://doi.org/10.1016/j.jrp.2005.08.002

Martingano AJ (2020) A Dual Process Model of Empathy (Doctoral dissertation).

Martingano AJ, Castano E (2020) Listening-in increases cognitive empathy relative to participating in a conversation. Manuscript submitted for publication.

Mascaro JS, Rilling JK, Tenzin Negi L, Raison CL (2013) Compassion meditation enhances empathic accuracy and related neural activity. Soc Cogn Affect Neurosci 8(1 Jan):48–55. https://doi.org/10.1093/scan/nss095

Miall DS, Kuiken D (1994) Foregrounding, defamiliarization, and affect: response to literary stories. Poetics 22(5):389–407

Morewedge CK, Preston J, Wegner DM (2007) Timescale bias in the attribution of mind. J Persona Soc Psychol 93(1 Jul):1. https://doi.org/10.1037/0022-3514.93.1.1

Ndalianis A (2007) Art cinema. The cinema book. pp. 83–87

Oatley K (2016) Fiction: simulation of social worlds. Trend Cogn Sci 20(8 Aug):618–628. https://doi.org/10.1016/j.tics.2016.06.002

Panero ME, Weisberg DS, Black J, Goldstein TR, Barnes JL, Brownell H, Winner, E (2016) Does reading a single passage of literary fiction really improve theory of mind? An attempt at replication. J Pers Soc Psychol 111(5), e46. https://doi.apa.org/doiLanding?doi=10.1037%2Fpspa0000064

Pino MC, Mazza M (2016) The use of "literary fiction" to promote mentalizing ability. PLoS ONE 11(8):e0160254. https://doi.org/10.1371/journal.pone.0160254

Plantinga C (1999) The scene of empathy and the human face on film. In:Plantinga Carl, Smith Greg M (ed.) Passionate views: film, cognition, and emotion. The Johns Hopkins University Press, Baltimore, pp. 239–55

Preacher KJ, Hayes AF (2004) SPSS and SAS procedures for estimating indirect effects in simple mediation models. Behav Res Methods Instrument Comput 36(4 Nov):717–731. https://doi.org/10.3758/bf03206553

Premack D, Woodruff G (1978) Does the chimpanzee have a theory of mind? Behav Brain Sci 1(4):515–526

Reeves B, Yeykelis L, Cummings JJ (2016) The use of media in media psychology. Media Psychol 19(1):49–71

Rooney B, Bálint KE (2018) Watching more closely: shot scale affects film viewers' theory of mind tendency but not ability. Front Psychol 8:2349

Rubio-Fernández P, Glucksberg S (2012) Reasoning about other people's beliefs: bilinguals have an advantage. J Exp Psychol 38(1):211

Santiesteban I, White S, Cook J, Gilbert SJ, Heyes C, Bird G (2012) Training social cognition: from imitation to theory of mind. Cognition 122(2):228–235. https://doi.org/10.1016/j.cognition.2011.11.004

Sarris A (2007) Notes on the 'Auteur' Theory in 1962. Kwartalnik Filmowy 59:6–17

Saxe R, Schulz LE, Jiang YV (2006) Reading minds versus following rules: dissociating theory of mind and executive control in the brain. Soc Neurosci 1(3-4):284–298

Schatz T (1981) Hollywood genres: Formulas, filmmaking, and the studio system. McGraw-Hill Humanities/Social Sciences/Languages

Schweinitz J (2011) Film and stereotype: A challenge for cinema and theory. Columbia University Press

Shaheen JG (2003) Reel bad Arabs: How Hollywood vilifies a people. Ann Am Acad Polit Soc Sci 588(1):171–193

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

Smith D, Schlaepfer P, Major K, Dyble M, Page AE, Thompson J, Chaudhary N, Salali GD, Mace R, Astete L, Ngales M, Vinicius L, Migliano AB (2017) Cooperation and the evolution of hunter-gatherer storytelling. Nat Commun 8(1 Dec):1–9. https://doi.org/10.1038/s41467-017-02036-8

Smith M (1995) Engaging characters: fiction, emotion, and the cinema. Clarendon Press, Oxford

Smith SM, McIntosh WD, Bazzini DG (1999) Are the beautiful good in Hollywood? An investigation of the beauty-and-goodness stereotype on film. Basic Appl Soc Psychol 21(1):69–80

Stein NL, Trabasso T (1985) The search after meaning: comprehension and comprehension monitoring. In Applied developmental psychology. Elsevier, pp. 33–58

Suh S, Trabasso TR (1993) Inferences during reading: converging evidence from discourse analysis, talk-aloud protocols, and recognition priming. J Memory Lang 32(3):279–300. https://doi.org/10.1006/jmla.1993.1015

Tan E (2013) The empathic animal meets the inquisitive animal in the cinema: notes on a psychocinematics of mind reading.

Tikka P, Kauttonen J, Hlushchuk Y (2018) Narrative comprehension beyond language: common brain networks activated by a movie and its script. PLoS ONE 13(7):e0200134. https://doi.org/10.1371/journal.pone.0200134

Tomasello M (2010) Origins of human communication. MIT press

Vaage MB (2010) Fiction film and the varieties of empathic engagement. Midwest Stud Philos 34:158–179

van Peer W (1986) Stylistics and psychology: investigations of foregrounding. Taylor & Francis

Visch V, Tan E (2008) Narrative versus style: effect of genre-typical events versus genre-typical filmic realizations on film viewers' genre recognition. Poetics 36 (4):301–315. https://doi.org/10.1016/j.poetic.2008.03.003

Wexler, B, Iseli M, Leon S, Zaggle W, Rush C, Goodman A, Imal AE, Bo E (2016) Cognitive priming and cognitive training: Immediate and far transfer to academic skills in children. Sci Rep 6(1). https://doi.org/10.1038/srep32859

Wiessner PW (2014) Embers of society: firelight talk among the Ju/'hoansi Bushmen. Proc Natl Acad Sci 111(39):14027–14035

Wollen P (1972) Signs and Meaning in the Cinema (Vol. 9). Indiana University Press

Young Cushman F, Hauser M, Saxe R (2007) The neural basis of the interaction between theory of mind and moral judgment. Proc Natl Acad Sci 104(20 May):8235–8240. https://doi.org/10.1073/pnas.0701408104

Young L, Camprodon JA, Hauser M, Pascual-Leone A, Saxe R (2010) Disruption of the right temporoparietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments. Proc Natl Acad Sci 107(15 APR):6753–6758. https://doi.org/10.1073/pnas.0914826107

Zunshine L (2007) Why Jane Austen was different, and why we may need cognitive science to see it. Style 41(3):275–298

Zunshine L (2019) What Mary Poppins Knew: theory of mind. Children's Literat Hist Narrative 27(1):1–29

Zwaan RA, Langston MC, Graesser AC (1995) The construction of situation models in narrative comprehension: An event-indexing model. Psychol Sci 6 (5):292–297

Competing interests
The author declares no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1057/s41599-021-00793-y.

Correspondence and requests for materials should be addressed to E.C.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.