

**An Investigation Into the Relationship Between Fiction and Nonfiction Reading Exposure, and
Factors of Critical Thinking**

Helena Hollis

UCL Department of Information Studies

ORCID: <https://orcid.org/0000-0003-1950-0657>

Please cite as:

Hollis, H. (in press) An investigation into the relationship between fiction and nonfiction reading exposure, and factors of critical thinking. *Scientific Study of Literature*.

The Scientific Study of Literature journal is available at: <https://benjamins.com/catalog/ssol>

Abstract

Reading fiction has been associated with improved social and imaginative reasoning that could lead to improved critical thinking. This observational study investigated the relationship between fiction and nonfiction exposure, narrative transportation, and factors of critical thinking (critical thinking disposition, and epistemological orientation). Self-selecting participants ($N = 335$) completed an online survey including an author recognition test and self-report scales. Fiction scores were significantly associated with higher critical thinking disposition, while nonfiction had an inverse effect correlating with lower disposition. Fiction reading was associated with decreased absolutism, and nonfiction score conversely with higher absolutism. Total and nonfiction print exposure were associated with lower multiplism, with no significant association for fiction. Total and fiction print exposure were associated with higher evaluativism, with no significant association for nonfiction. Narrative transportation mediated some of these relationships. These findings provide a basis for further research into reading fiction and nonfiction, and critical thinking.

Keywords: fiction; print exposure; narrative transportation; critical thinking; epistemological orientation; critical thinking disposition.

An Investigation into the Relationship Between Fiction and Nonfiction Reading Exposure, and Factors of Critical Thinking

Fiction reading has been linked to a variety of social and cognitive outcomes, and as such arguments have been made for engagement with fiction acting as a simulative training activity (Oatley, 2011). Valuing fiction reading as a source of these gains fits into a wider context of argument in favour of the value of the arts, in which interpersonal skill and critical thinking (CT) are at the foreground as stipulated benefits (Dumitru, 2019). However, the relationship between fiction reading and CT is an under-researched area. This study therefore aimed to investigate how individual differences in reading, specifically fiction and nonfiction print exposure, impact individuals' CT disposition and their epistemological orientation (EO). These outcome variables have been selected as key facets of CT. The study also investigated the role of narrative transportation (NT) as a mediator between reading and its outcomes. This study was conducted as part of a wider doctoral research project investigating the effects of fiction reading on CT.

Reading and Critical Thinking

CT is argued to be a vital capacity for social and democratic engagement (Holma, 2015). This has especially been emphasised in light of current 'fake news' concerns, which highlight the need for individuals to evaluate each piece of information they encounter as heuristics of source quality have been shown unreliable (Batchelor, 2017; Lutzke et al., 2019). Indeed, CT is often considered an essential component of literacies needed to tackle the ways that fake news and related issues of media trustworthiness can threaten democracy (Mason et al., 2018), emphasising the role of CT in social democratic engagement. Subsequently it has been argued that CT is a primary enabling force for change, as all people of all genders, ethnicities, and backgrounds can use their capacity to think critically and thus change and shape society (hooks, 1994). As such, modern CT conceptions extend beyond logical or syllogistic reasoning, encompassing empathy and social skill (Ennis, 2016). CT has been characterised as a dialogic process, as critical thinkers must generate arguments, but also anticipate rebuttals, and thus imaginatively model a back-and-forth dialogue as part of the thinking

process (Kuhn, 2019; Mercier & Sperber, 2011). Therefore, the ability to imaginatively model the minds of others, and engage in their potential trails of thought, it is an essential activity of CT (Paul, 1987); this also entails empathy (Thayer-Bacon, 2000), and openness (hooks, 1994). Furthermore, CT is also inherently reflective, as in addition to imagining the minds of others, critical thinkers must also continuously question the workings of their own mind (Weinstock et al., 2017). Paul (1987) argues that this broader conception of CT that encompasses imaginative, simulative, and social facets is CT in the “strong sense” while purely logic-focused conceptions capture CT only in the “weak sense.” For the purposes of this study, “strong sense” CT was the conception of interest, and it is in its strong sense that CT may be linked with reading.

Empathy is an important aspect of CT in the strong sense; however, there is a lack of definition of empathy given in the literature on CT. More broadly, approaches to empathy from psychology and philosophy can be used to define it and link it to both CT and reading. Empathy can be described as the ability to understand another’s emotions and feel those emotions alongside them, with different types of empathy arising at different levels in different scenarios for different individuals (Cuff et al., 2014). Richmond (2004) identifies the origin of the concept of empathy in the field of aesthetics as an “imaginative reconstruction,” which positions empathy as an act of simulation. This simulationist position is also used in research on fiction and its influence on empathy (e.g., Oatley, 2011). Empathy is furthermore dependent on other factors, such as memory, and self-reflection (Decety & Jackson, 2004), as one must be able to recall past emotional experiences to match them to the observed person. As a bridge to the experiences of others and for self-reflection, empathy has been argued to be a component of or requisite for CT (e.g., Mulnix, 2012; Paul, 1987; Thayer-Bacon, 2000). For the purposes of this study, the view of empathy as an integral component of CT is adopted (Thayer-Bacon, 2000).

Thayer-Bacon perhaps most thoroughly argues for the inclusion of empathy within understandings of CT, as “receiving others, feeling with others” (Thayer-Bacon, 1992, p. 10), which permits us to assess the knowledge of others as well as their motives, thus permitting thorough

engagement in thinking critically about their positions. However, experimental findings testing this hypothesis are somewhat scarce, with only small studies showing positive correlation between CT disposition and empathy (Ekinci & Aybek, 2010; Jeong, 2015). Conversely, it should be noted that arguments on empathy and CT can also be made in the opposite direction; Paul Bloom (2016) has argued that empathy in fact impairs moral decision making by clouding rational judgement and making us more susceptible to bias. However, Bloom's argument rests on a narrow definition of empathy that emphasises emotional contagion and does not take account of the above discussed facets of empathy. Overall, there is some indication that empathy may play a role in CT, but further research is needed.

Decety and Jackson (2004) note that empathy is highly flexible and can be trained and improved through intervention. Fiction reading has been found to be associated with increased empathy and is often described as an efficacious intervention for training empathy (e.g., Bal & Veltkamp, 2013; Mar et al., 2006; Stansfield & Bunce, 2014). If fiction reading improves empathy, and empathy is necessary for CT, then an argument for fiction reading leading to improvements in CT can be supported. It has also been argued that using empathy-inducing materials in class may help CT by preventing an 'us versus them' perspective on controversial issues (DioGuardi, 2016); this resonates with research showing reading about the lives of others can help break down cultural barriers (Hakemulder, 2001). There is therefore some indication that engagement with fiction may yield CT improvement via empathy.

Existing research into the effects of reading, especially fiction, has found increases in factors that may lead to increased CT ability. Readers have been found to make evaluations throughout the fiction reading experience, considering issues raised from different perspectives, and reflecting upon their own lives and beliefs (Özyürek & Trabasso, 1997); thus CT is directly involved and practiced when reading. Furthermore, it has been theorised that fiction reading drives increases in theory of mind (Zunshine, 2006), and this has been empirically supported (e.g., Kidd & Castano, 2013), with broad support for an association between reading fiction and increased social skill, theory of mind

and empathy (Mumper & Gerrig, 2017; Tamir et al., 2016). Authors argue this benefit stems from an imaginative modelling of different perspectives as one engages with the minds of fictional characters, a depiction that strongly resonates with dialogic conceptualisation of CT. Reading is thought to impact empathy due to the way the narrative transports the reader (Johnson, 2012), and its broader impacts are often attributed to the training of simulative capacities (Oatley, 2011). It therefore may be that by reading fiction we are building empathetic, imaginative capacities which in turn can be deployed in critical thought.

In addition to imagining the mental and emotional states of others, further imaginative modelling of divergent possibilities is also needed for CT. This fits under the umbrella of counterfactual thinking, in which alternative possibilities are generated in order to enable critical evaluation of the past or present and to speculate about the future (Byrne, 2016). Corballis (2009) points out that our ability to share our episodic memories with one another through language permits us to share the material from which we can construct our counterfactual imaginings. As fictional material is also incorporated into our episodic memory (Zunshine, 2006), fiction therefore also builds up the store of content which can then be re-deployed counterfactually. Empirical evidence supports this argument, as reading fiction has related to greater openness in imagining alternative possibilities, or counterfactual reasoning (Black et al., 2017). This again resonates with the need for openness and counterfactual imagination in CT.

This is not to say nonfiction reading plays no role, as it is associated with and valued for “informational reading” providing a means of knowledge building (Hampson Lundh et al., 2018), and a knowledge base is also a requirement for CT (Byrnes & Dunbar, 2014). However, fiction may play a special role in building CT capacity via its offline and safe processing, which has been posited to bypass our defences (Bal & Veltkamp, 2013; Djikic et al., 2009). The offline safety of fictional experiences has been likened to a flight simulator, as a means of testing and learning social skills, rather than piloting, without real-world consequences (Oatley, 2011). Furthermore, fiction has been found to circumvent defences we may have around certain themes or subject areas, which we might

avoid in nonfiction, as the themes and characters we willingly engage with in fiction reading seem more diverse (Hakemulder, 2008), and riskier (Zunshine, 2006), than those we typically encounter in reality, and thus may further help build openness which is essential to CT. For example, reading Bret Eaton Ellis' *American Psycho* may be highly uncomfortable, yet many readers are willing to engage with the fictional narrative; unlike nonfictional accounts of analogous violence, the novel has been argued to provide "a way of experiencing violence without censorship, but also without being overwhelmed" (Serpell, 2010, p. 67). This, it is argued, in fact permits for a deeper critique of violence, and by extension other themes we may be closed off to in nonfictional renderings could be opened for inspection in fictional accounts without our censoring them in order to protect ourselves from becoming overwhelmed. Indeed, art more broadly has been argued to uniquely accommodate negative emotion such that it is safely distanced and can ultimately be embraced in deeper processing (Menninghaus et al., 2017).

This study is therefore based upon the hypothesis that fiction can increase CT directly, by training our simulative capabilities that underlie CT, and indirectly by increasing our theory of mind, empathy, counterfactual reasoning, and self-reflective capacity, all of which in turn can be deployed in CT. While nonfiction reading can also lead to CT gains through knowledge building, knowledge only plays a small role in modern concepts of CT, with greater weighting given to interpersonal and counterfactual imagining and reasoning (e.g., Ennis, 2016). The gains from fiction reading are expected to be greater as they stem from a more diverse array of factors, more closely aligned with CT in the "strong sense" (Paul, 1987). However, the assessment of CT is a contested topic with a multiplicity of approaches, and tests of CT require significant participant time commitments (Abrami et al., 2008). Therefore, as a preliminary correlational exploration of a potential relationship, this study sought to test these suppositions about reading relating to CT by establishing whether there is an association with reading engagement and factors associated with CT. As such, CT itself was not measured but CT disposition and EO were taken as its indicators.

Narrative Transportation and Critical Thinking

Being transported by a narrative is the experience of being immersed into the story world such that it is vividly felt, and perceived (Gerrig, 1993) NT is argued to be an active process, and as such some level of effort is required to pay attention and focus on the narrative in order to be transported (Green & Brock, 2013). NT can have trait and state manifestations, with trait NT sometimes termed “transportability” (Dal Cin et al., 2004; Mazzocco et al., 2010); for purposes of this study, the term “NT” will be used for simplicity, though measuring trait transportability. In addition, empathy has also been argued to be a requirement for NT (van Laer et al., 2013); however, as NT has been found to heighten empathy in reading (Johnson, 2012), this relationship may best be thought of as reciprocal. Busselle and Bilandzic (2008) argue NT takes place in the construction of mental models of the narrative, which unifies conceptions of narrative transportation with those of fiction reading as mental modelling (Oatley, 2011). Finally, retrospective reflection has been found to be a mediator between narrative transportation and belief and attitude outcomes, as the ability to link aspects of the narrative to one’s past experiences facilitates the model building process (Hamby et al., 2017). Thus, NT has been associated with many of the factors of modern CT conceptions.

NT has furthermore been shown to have powerful effects on wide ranging types of attitudes and decision making (Appel & Richter, 2007; Hamby et al., 2017). However, findings on the relationship between NT and CT are mixed. Being highly transported by an advertising narrative has been shown to reduce CT while increasing affect (Escalas, 2004). Furthermore, a lower tendency to think of counterarguments or to be resistant towards what is being presented when highly transported in a narrative has been shown (Dal Cin et al., 2004; Green & Brock, 2013). These findings have led many to hypothesise that NT reduces CT (Busselle & Bilandzic, 2008). However, Hoeken and Fikkers (2014) dispute this and argue that the relationship between NT and CT is in fact far more complex, as they found participants displayed high levels of issue-relevant evaluative thinking while transported by a narrative in which a protagonist or antagonist presented a persuasive argument. It

should be noted that these studies all investigated the effects of state NT by a given narrative upon attitudes towards the topic of that narrative. They did not investigate a relationship between trait NT and trait CT dispositions and abilities. This is an area that requires further study. Furthermore, the prior research on NT and thinking critically has focused heavily on narratives written for persuasiveness (Hamby et al., 2017); this is a distinctive type of narrative, which seems very different from the novels and short stories referred to in studies on fiction and its impacts on social and cognitive abilities. This presents a research gap, where different types of narrative could be tested and different effects of NT upon critical thought could be found.

Critical Thinking Disposition

CT disposition is an essential component of CT (Ennis, 2016; Facione, 1990). It is defined as “a constellation of attitudes, intellectual virtues or habits of mind” capturing one’s willingness to engage in CT (Sosu, 2013, p. 108), which implies at least some ability to do so. Intellectual virtues can be understood with different philosophical framings but broadly entail characteristics that we can develop and utilise to underpin our knowledge, such as openmindedness, or thoroughness (Greco, 2000). Thus, to be disposed to think critically does not simply mean having a desire to do so; it also encapsulates underpinning characteristics needed to do so, meaning CT disposition is more than motivation alone. For example, being very thorough and systematic when approaching information is an element of CT disposition, implying both motivation in taking the time and effort, and some ability to examine information step by step. This distinguishes the way authors conceptualise CT disposition from other psychological models of motivation (e.g., Cacioppo & Petty, 1984). As CT is effortful (Byrnes & Dunbar, 2014), the disposition towards it has been argued as a necessary prerequisite to deploying the required effort and engaging in CT practice (Weinstock et al., 2017). CT disposition has been shown to correlate positively with performance on problem-based learning tasks (Pu et al., 2019) and with critical thinking skill (Colucciello, 1997). Furthermore, CT disposition has been correlated with social emotional learning (Arslan & Demirtas, 2016), which entails empathy and broader social skills as part of CT. Engagement with a digital storytelling activity has been shown

to increase CT disposition, which the author argues to be due to the dialogic nature of the activity and its promotion of self-reflection (Chan, 2019); this gives an indication that engagement with narrative may be a relevant driver for CT disposition.

Epistemological Orientation

Epistemological beliefs are those pertaining to what knowledge is and how it can be gained, and these have also been identified as an important aspect of CT (Kuhn, 1991). These beliefs can group into epistemological orientation (EO), typically divided into three categories: (a) 'absolutist,' entailing a belief that statements are either true or false; (b) 'multiplist,' entailing a belief that statements reflect an entirely subjective reality; and (c) 'evaluativist,' entailing a belief that statements can be judged based on some evidence and criteria (Hofer & Pintrich, 1997; Kuhn et al., 2000). Experimental findings support the supposition that EO is an important contributing factor for CT (Hyytinen et al., 2014; King & Kitchener, 2004; Kuhn, 1991; Liu et al., 2011; Tsui, 2000).

Evaluativist EO is considered most conducive of CT (Kuhn et al., 2000), as one must believe that statements can be effectively judged in order to attempt to do so. This evaluative stance has been cast as 'epistemic vigilance' as it necessitates being on guard for both absolutist and multiplist assumptions that could curtail CT (Mercier & Sperber, 2011). Furthermore, this stance is inherently reflective, as this epistemic vigilance is applied to one's own approaches, and this further connects it to the self-reflective nature of CT (Weinstock et al., 2017). In addition to being linked with CT, EO has been associated with different reading approaches and comprehension (Strømsø et al., 2008). The possible influence of reading upon EO can be inferred from studies showing fiction reading association with more open perspectives; for example, Fong et al. (2015) found an increased gender role egalitarianism in fiction readers, suggesting openness. Furthermore, reading short stories has been found to reduce a need for cognitive closure (Djikic et al., 2013), which also implies openness. However, while these studies may be suggestive of lower absolutism in fiction readers, this does not necessarily mean increased evaluativism, and EO itself has not been tested. This is therefore an area for further investigation.

Correlating Fiction and Nonfiction Reading With CT Disposition and EO

There is a surprising dearth of research into individual differences in reading, particularly fiction, and CT. In a broader sense, engagement with arts and humanities is commonly argued to improve CT (Dumitru, 2019), but more fine-grained analysis of specific arts and their role in CT development beyond academic settings is lacking. The existing research on reading and CT tends to focus on the link with reading comprehension (Aloqaili, 2012), rather than any influence of reading material on thinking. Studies that have addressed CT gains through reading have made this connection through critical reading, treating the text as a prompt for CT, thus providing direct training for CT (Shihab, 2011; Tabačková, 2015). For instance, developing literary understanding as part of the study of literature has found to be mediated by disposition to think critically and moderated by CT skills (Koek et al., 2016). Such approaches focus on the use of literary texts in classroom contexts with instruction, rather than on unguided reading, or on fiction beyond what is deemed literary. More broadly, reading voluntarily, particularly reading books, has been associated with increased CT, but without distinguishing fiction from nonfiction (Hawkins, 2012). Where fiction reading has been associated with CT, this tends to be about the specific topics contained within the fictional text; for example, science fiction can be treated as a means of conceptualising the future in order to think critically about it, to the extent that has been treated as a research method in AI forecasting (Avin, 2019). Others have suggested ways fiction could be used in teaching CT but without testing the posited benefits (Pardede, 2019). There is some anecdotal suggestion for the CT benefits of fiction reading; Oatley (2011) notes a case of the *Literature for All of Us* reading group reporting increased critical thinking and problem solving in their participants (p. 186). However, at the time of this writing, to the author's knowledge, no research has been conducted to ascertain a correlation between fiction reading and CT or associated factors.

By assessing two outcome variables associated with CT, this study aims to establish whether there is any correlation between reading fiction and nonfiction, EO, and CT disposition, as an indication of a potentially fruitful direction of research on reading and CT.

Hypotheses:

H1: print exposure will correlate with stronger critical thinking disposition, and with more sophisticated epistemological orientation.

H2: fiction reading will correlate more strongly than nonfiction reading with stronger critical thinking disposition, and with more sophisticated epistemological orientation.

H3: narrative transportation will mediate the relationships predicted in H1 and H2.

Method

Participants

Adult (18 years of age or above) participants were recruited using online calls for participation on social media (twitter and reddit), as well as online study promotion websites (callforparticipants.com, Survey Circle, Survey Swap). The study was also promoted in London public spaces using posters and bookmarks. Participants self-identified as being fully proficient in English. A £20 voucher was offered as a prize for 10 randomly selected participants. A total of 353 responses were obtained. From these, 18 were removed for incompleteness, yielding 335 responses.

Participants included in both the CT disposition and EO analyses had a mean age of 33, median 30, with the youngest being 18 and oldest 77.

Participants who completed both scales included 224 females and 87 males; participants who completed the CTDS only and were not included in the EO analyses included 216 females and 86 males. In both cases, this means that 72% of the sample was female, and 28% was male.

Highest obtained educational level was coded into the following categories: GCSE or equivalent and vocational: 8% (CTDS $n = 25$, EO $n = 25$); A level: 7% (CTDS $n = 22$, EO $n = 21$); bachelor's degree: 34% (CTDS $n = 107$, EO $n = 103$); master's degree: 42% (CTDS $n = 128$, EO $n = 126$); doctoral: 9% (CTDS $n = 29$, EO $n = 27$).

Measures

Demographic information (age, gender, educational level) was collected using open questions, and responses for gender and education were then coded into categories.

The ART-G was used as a measure of print exposure (Mar & Rain, 2015). This is checklist of 200 author names (e.g., "Albert Camus"), including 40 foil items (i.e., names that do not belong to authors, e.g., "Frank Bluth"). Participants are asked to check those names in the list which they recognise to be authors. The scale is scored as a numerical variable, yielding total numbers of correctly identified authors minus foils checked, and scores for fiction (110 items in total) and nonfiction (50 items in total). The scale has been found to have good reliability for fiction (Cronbach's $\alpha = .90$) and nonfiction (Cronbach's $\alpha = .82$).

The Fantasy subscale of the Interpersonal Reactivity Index (IRI-F) was used as a measure of Narrative Transportation (NT) (Davis, 1983). While other measures of NT are available, this scale was selected for its brevity so as to minimise participant burden and furthermore to be in keeping with studies used to formulate the hypotheses of this study (e.g., Mar et al., 2009). The scale was used to measure NT in line with other similar research that omitted one item of the scale which does not pertain to immersion in a narrative, and the removal of which has been found to increase its internal reliability (Cronbach's $\alpha = .81$) (Mar et al., 2009). This resulted in a six- item Likert scale, with two items reverse coded, with five options of agreement rating from strongly disagree to strongly agree. An example item included, "I really get involved with the feelings of the characters in a novel." No instruction for scoring is present in the manual (Davis, 1983); therefore, the IRI-F was scored as a numerical variable in keeping with previous studies (e.g., Mar et al., 2009).

The Critical Thinking Disposition Scale (CTDS) was used as a measure of CT disposition (Sosu, 2013). This is a Likert scale with 1-5 agreement rating, containing 11 items. Items include those that capture habits (e.g., "I usually check the credibility of the source of information before making judgements") and those that touch on underlying values that could be deemed intellectual virtues (

e.g., "It's important to understand other people's viewpoint on an issue"). The scale has been shown to have high internal reliability (Cronbach's $\alpha = .81$). It was scored as a numerical variable.

The Justifying Conclusions Inventory (JCI) was used as a measure of EO (McGinnis, 2016). This is a Likert scale with 1-6 agreement rating, containing 23 items. The items represent three categories of EO: absolutism (e.g., "For most debatable issues, there is a right position and a wrong position"); multiplism, (e.g., "Deciding among conclusions is unnecessary because everyone has a right to their opinion"); evaluativism, e.g., "Understanding debatable issues is an ongoing process requiring the evaluation of new evidence"). Scoring is performed using mean responses for each category of EO. The JCI has good reliability for each subscale, with Cronbach's alpha values of absolutism $\alpha = .86$; multiplism $\alpha = .77$; evaluativism $\alpha = .78$.

Procedure

Participants followed a link to the online survey, where they were presented with the information sheet and consent form. Upon consenting, participants were asked to complete demographic information. They were then given the scales in the following order: IRI-F, ART-G, CTDS, and JCI. Finally, participants were offered the option of signing up for the prize draw and debriefed. All answers were anonymous.

Participants were excluded based on how many foil items in the ART-G they checked so as to remove responses with excessive guessing. The cut-off for foil checking was set at one SD above the mean; the mean number of foils checked was 0.4, with a SD of 1.22; therefore, two items was the maximum permitted number of foils checked. Based on this, 11 participants were excluded. A further three participants were removed as they did not select any items in the ART-G. This decision was taken to avoid non-genuine answers; as a prize draw was offered, there was some incentive to click-through the questionnaire, and the ART-G was the only scale where no selection was mandatory. Given the high profile of several names in the list, a zero-recognition rate was deemed highly unlikely. This yielded 321 responses.

Furthermore, 11 participants did not complete the JCI, which was presented last in the order of items in the survey, though they did complete all other parts of the survey. These responses were kept for analysis with CT disposition as the outcome variable but excluded for analysis of EO as the outcome variable.

Finally, participants who did not disclose their gender ($n = 9$) or identified as neither male nor female ($n = 1$) were excluded in models with gender as a variable. This is due to the low sample size these gender categories would contain if included in the analysis.

The significance threshold was set at .05. Where variance was equal between groups, even in the event of non-normal distributions, ANOVA was used due to robust sample size (Blanca et al., 2017). Analysis was conducted using R 4.0.0. in R Studio version 1.2.5042.

Results

Descriptive Statistics

The ART-G total score is the total number of correct items selected, minus the number of foils selected. The mean total score was 30.62 (SD 22.78), with a median of 26. The minimum total score was 1, and the maximum 116. A Shapiro-Wilk normality test found ART-G total scores not to be normally distributed ($W = .92, p < .001$).

The mean number of ART-G fiction items checked was 23.57 (SD 18.57), with a median of 19, and responses ranging from 0 to 99; the mean response therefore represents a 22% correct fiction recognition rate. A Shapiro-Wilk normality test found ART-G fiction scores not to be normally distributed ($W = .91, p < .001$).

The mean number of nonfiction items checked was 7.26 (SD 7.29), with a median of 5, and responses ranging from 0 to 45; the mean response represents a 14% correct nonfiction recognition rate. A Shapiro-Wilk normality test found ART-G nonfiction scores not to be normally distributed ($W = .78, p < .001$).

Figures 1-3 show ART-G score distributions.

Figure 1: Distribution of ART-G total scores

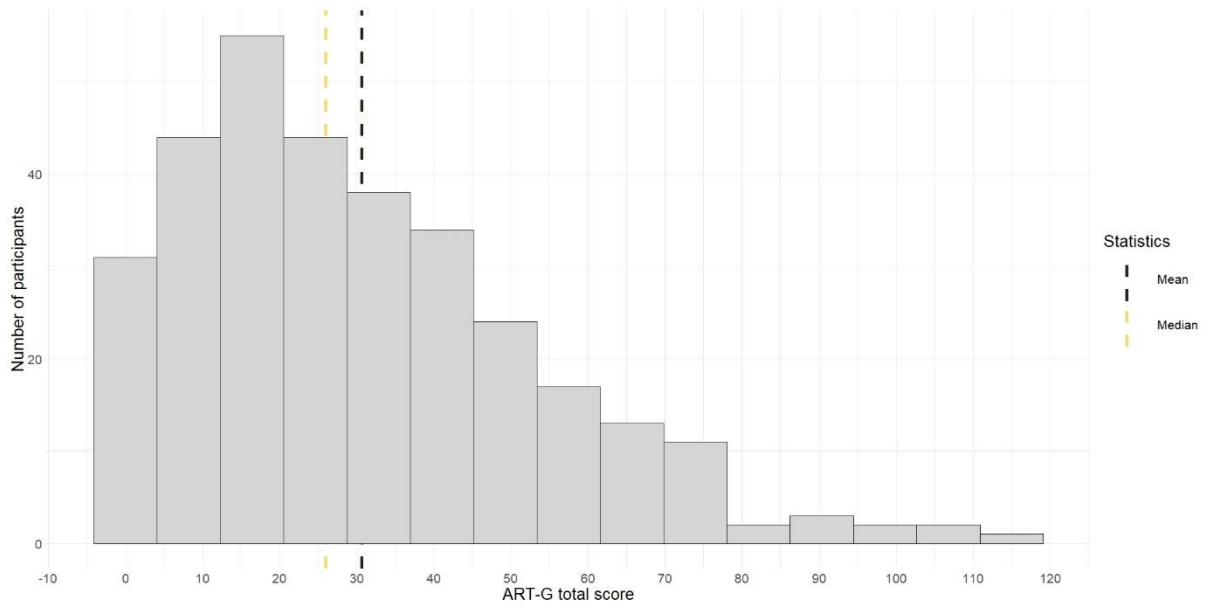


Figure 2: Distribution of ART-G fiction scores

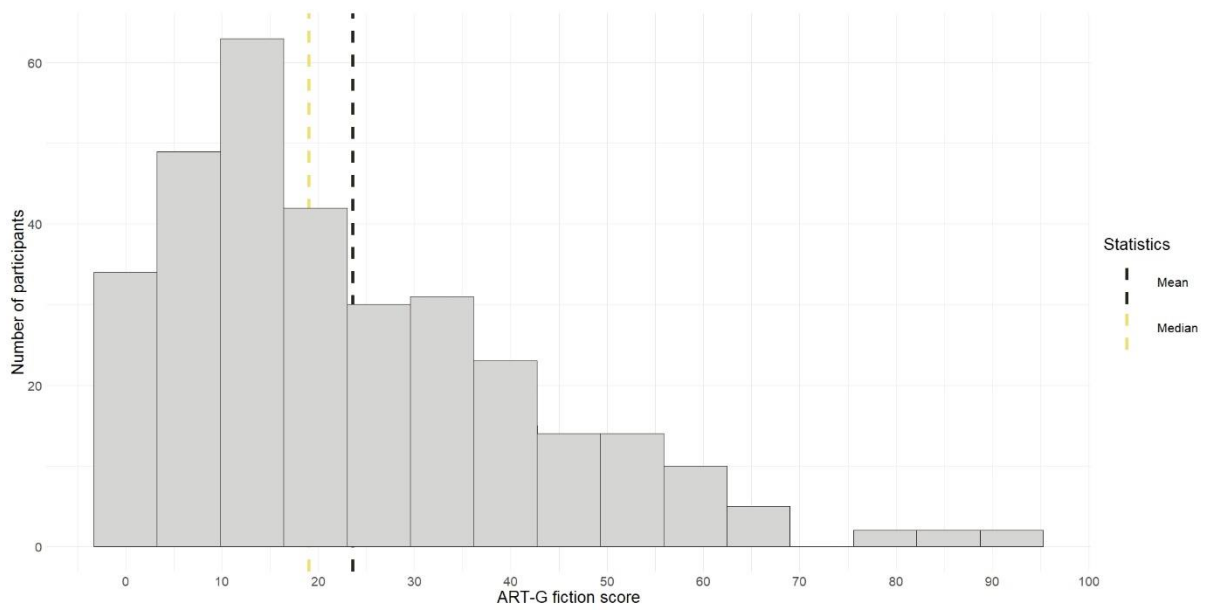
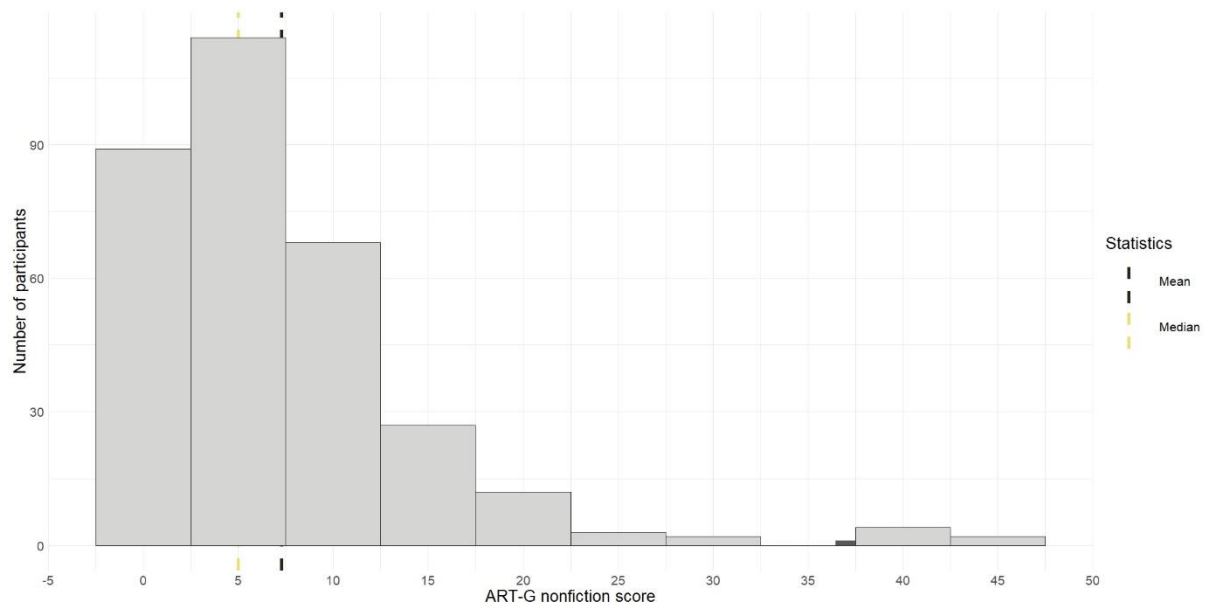


Figure 3: Distribution of ART-G nonfiction scores



Participants' mean scores on the IRI-F were used. The grand mean score on the IRI-F was 4.02, with a *SD* of 0.8. A Shapiro-Wilk normality test found IRI-F scores not to be normally distributed ($W = .96, p < .001$).

Participants' mean scores on the CTDS were used. The CTDS had a grand mean of 4.11, *SD* 0.54. A Shapiro-Wilk normality test found CTDS scores not to be normally distributed ($W = .93, p < .001$).

Participants' mean JCI scores on each scale were as follows: multiplism $M = 3.12, SD = 0.99$, range 1 – 6; evaluativism $M = 5.09, SD = 0.65$, range 2.71 – 6; absolutism $M = 2.62, SD = 1.16$, range 1 – 6. A Shapiro-Wilk normality test found JCI scores not to be normally distributed: multiplism ($W = .99, p = .02$); evaluativism ($W = .94, p < .001$); absolutism ($W = .93, p < .001$).

Modelling CT disposition

CT Disposition Model Selection

In order to develop a model to test with CT disposition as the outcome variable, relationships between this and the measured demographic variables (age, education, and gender) were tested to identify which of these ought to be included.

Spearman's rank correlation found no significant association between CTDS scores and age ($\rho = .10, p = .061$).

An ANOVA showed a significant difference in CTDS scores across educational levels [$F(4, 316) = 3.72, p = .006$]. Tukey's multiple comparisons of means showed the only significant differences to be between GCSE or Equivalent and Vocational and Master's degree with an average increase of 0.33 (95% CI 0.02, 0.64, $p = .035$). Educational level was therefore considered a relevant predictor variable for CT disposition.

An ANOVA was used and showed no significant difference in CTDS scores across genders [$F(3, 317) = 0.49, p = .698$]. Gender was therefore not considered a relevant predictor variable for CT disposition.

Next, the direct relationships between the predictor variables of interest (ART-G total, fiction and nonfiction scores; NT level) and CT disposition level were assessed. Spearman's rank correlation found significant relationships for all: ART-G total score ($\rho = .14, p = .014$); ART-G fiction score ($\rho = .18, p = .001$); ART-G nonfiction score ($\rho = .16, p = .004$); IRI-F score ($\rho = .33, p < .001$). It is notable that of the ART-G scales, fiction had the strongest correlation with CTDS score. However, IRI-F score was more strongly correlated with CTDS, suggesting NT has a stronger association with CT disposition than print exposure.

Finally, IRI-F scores were tested for correlations with the ART-G scales. Spearman's rank correlation found no significant relationship with total scores ($\rho = .07, p = .214$); a significant relationship with fiction ($\rho = .15, p = .006$); no significant relationship with nonfiction ($\rho = -.06, p = .301$). This suggests that increased NT is only associated with greater fiction exposure.

Based on these individual analyses, the variables of interest in predicting CT disposition level were education level, NT, ART-G total, fiction and nonfiction scores. These variables were therefore used in a hierarchical linear regression. Education was entered into the regression first so as to control for its effects. Due to the conceptual framing of NT in relation to reading and its outcomes, NT was tested as a mediator. However, as this only correlated with ART-G fiction scores, it was only

expected to mediate this relationship. As ART-G total scores are the sum of fiction and nonfiction, these were treated separately so as to avoid collinearity.

CT Disposition Model Results

Table 1: Hierarchical regression results for CTD and ART-G total score

| Predictor Variable | B(SE) | 95% CI for B | | β | R^2 | ΔR^2 |
|--------------------|-------------|--------------|------|----------|-------|--------------|
| | | LL | UL | | | |
| Step 1 | | | | | | |
| Constant | 3.74 (0.10) | 3.54 | 3.94 | <.001*** | .04 | .04 |
| Educational level | 0.11 (0.03) | 0.05 | 0.17 | <.001*** | | |
| Step 2 | | | | | | |
| Constant | 3.68 (0.11) | 3.48 | 3.89 | <.001*** | .05 | .01 |
| Educational level | 0.10 (0.03) | 0.05 | 0.16 | <.001*** | | |
| Total reading | 0.00 (0.00) | 0.00 | 0.00 | .075 | | |

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

Educational level was found to have as significant relationship with CT disposition. When ART-G total scores were added to the regression, no significant relationship was shown, and there was no significant improvement in model fit. Therefore, when educational level is controlled for, total print exposure is not associated with increased CT disposition.

Next, the hierarchical regression process was repeated, with nonfiction and fiction differentiated. Nonfiction was entered into the regression first, as fiction is the predictor variable of primary interest in this study and therefore was treated with the most conservative approach.

Table 2: Hierarchical regression results for CTD and ART-G fiction and nonfiction score

| Predictor Variable | B(SE) | 95% CI for B | | β | R^2 | ΔR^2 |
|--------------------|-------------|--------------|------|----------|-------|--------------|
| | | LL | UL | | | |
| Step 1 | | | | | | |
| Constant | 3.74 (0.10) | 3.54 | 3.94 | <.001*** | .04 | .04 |
| Educational level | 0.11 (0.03) | 0.05 | 0.17 | <.001*** | | |
| Step 2 | | | | | | |
| Constant | 3.78 (0.10) | 3.58 | 3.99 | <.001*** | .05 | .01 |

| | | | | | | |
|--------------------|--------------|-------|-------|----------|-----|--------|
| Educational level | 0.11 (0.03) | 0.06 | 0.17 | <.001*** | | |
| Nonfiction reading | -0.01 (0.00) | -0.02 | 0.00 | .064 | | |
| Step 3 | | | | | .11 | .06*** |
| Constant | 3.70 (0.10) | 3.50 | 3.90 | <.001*** | | |
| Educational level | 0.10 (0.03) | 0.05 | 0.16 | <.001*** | | |
| Nonfiction reading | -0.02 (0.00) | -0.03 | -0.01 | <.001*** | | |
| Fiction reading | 0.01 (0.00) | 0.00 | 0.01 | <.001*** | | |

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

As was the case with ART-G total scores, adding nonfiction scores in step two found no significant relationship and did not significantly improve model fit. However, when fiction scores were added in step three, model fit was significantly improved. This model tells us that when fiction and nonfiction exposure are accounted for, educational level remains a significant predictor of CT disposition. Additionally, when educational level and fiction exposure are accounted for, nonfiction exposure in fact has a significant negative association with CT disposition, with mean CTDS score decreasing on average by -0.02 per additional nonfiction author recognised. Finally, when educational level and nonfiction exposure are controlled for, fiction exposure had a significant positive association with CT disposition, and as each additional fiction author was recognised mean CTDS score increased on average by a small amount (0.01).

Mediation by NT Upon CT Disposition

Model-based causal mediation analysis was performed (using the 'mediation' package for R), with nonparametric bootstrapped (1000 simulations) confidence level estimation using the percentile method, to calculate the average causal mediation effect (Tingley et al., 2014). Mediation by NT was not tested for total print exposure, as this was not found to have a significant association with CT disposition. NT mediation was tested for both the nonfiction and fiction relationships with CT disposition.

Results from the causal mediation analysis showed that nonfiction scores had an insignificant direct effect of 0.00 (95% CI -0.09, 0.01; $p = .962$), a significant indirect effect via NT of -

0.01 (95%CI -0.01, 0.00; $p = .01$), and an insignificant total effect of -0.01 (95% CI -0.01, 0.01; $p = .38$).

Therefore, NT was shown to fully mediate the relationship between nonfiction and CT disposition.

Results from the causal mediation analysis showed that fiction scores had a significant direct effect of 0.01 (95% CI 0.01, 0.01; $p = .01$), a significant indirect effect via NT of 0.02 (95% CI 0.01, 0.02; $p < .001$), and a significant total effect of 0.01 (95% CI 0.01, 0.01; $p < .001$). The proportion of the total effect mediated by NT was 0.37 (95% CI 0.14, 0.74). Therefore, NT was shown to mediate the relationship between fiction and CT disposition.

Modelling EO

EO Model Selection

The demographic variables were first tested for associations with different EO types, so as to ascertain which variables ought to be included in regression models as controls.

Spearman's rank correlation found no significant association between JCI multiplism scores and age ($\rho = .09$, $p = .107$); nor absolutism ($\rho = -.11$, $p = .051$); however, increased evaluativism was associated with greater age ($\rho = .17$, $p = .002$). Therefore, age was considered to be a relevant control variable for evaluativism.

An ANOVA showed a significant difference in JCI multiplism scores across educational levels [$F(4, 306) = 2.45$, $p = .046$]. Tukey's multiple comparisons of means showed the only significant differences to be between Bachelors' degree and Doctoral levels, with an average decrease of -0.61 (95% CI -1.19, -0.03, $p = .035$). Educational level was therefore considered a relevant predictor variable for multiplist EO. Next, an ANOVA showed a significant difference in JCI evaluativism scores across educational levels [$F(4, 306) = 2.74$, $p = .029$]. Tukey's multiple comparisons of means showed the only significant differences to be between A Level or equivalent and Doctoral levels, with an average decrease of 0.51 (95% CI 0.00, 1.01, $p = .049$). Finally, an ANOVA showed no significant difference in JCI absolutism scores across educational levels [$F(4, 306) = 0.99$, $p = .411$]. This suggests that educational level is a relevant control variable for multiplist and evaluativist EO.

Due to the very small sample size of the non-binary and not stated gender categories, these were removed for the purposes of comparing EO scores across genders, and only male and female participants were included ($n = 302$). Wilcoxon rank sum test with continuity correction showed a significant difference between genders for multiplism scores ($p < .001$) but none for evaluativism ($p = .481$) nor absolutism ($p = .294$). Therefore, gender was found to be a relevant control variable for multiplism.

Next, the direct relationships between the predictor variables of interest (ART-G total, fiction and nonfiction scores; NT level) and EO levels were assessed. For multiplism, Spearman’s rank correlation found significant negative relationships for all reading scores: ART-G total score ($\rho = -.23, p < .001$); ART-G fiction score ($\rho = -.13, p = .015$); ART-G nonfiction score ($\rho = -.33, p < .001$). IRI-F score had a significant positive relationship ($\rho = .11, p = .037$). For evaluativism, Spearman’s rank correlation found significant positive relationships for all reading scores: ART-G total score ($\rho = .25, p < .001$); ART-G fiction score ($\rho = .27, p < .001$); ART-G nonfiction score ($\rho = .23, p < .001$). IRI-F score also had a significant positive relationship ($\rho = .20, p < .001$). For absolutism, Spearman’s rank correlation found significant negative relationships for all reading scores: ART-G total score ($\rho = -.15, p = .001$); ART-G fiction score ($\rho = -.20, p < .001$); ART-G nonfiction score ($\rho = -.12, p = .028$). IRI-F score also had a significant negative relationship ($\rho = -.16, p = .003$). This suggests that all types of print exposure were associated with lower multiplism and absolutism, and higher evaluativism. Higher NT was associated with increased multiplism and evaluativism, and decreased absolutism.

EO Model Results

Absolutism. As no demographic variables were found to have a significant relationship with absolutism, these were not controlled for. Firstly, ART-G total scores were tested in a linear regression.

Table 3: Hierarchical regression results for absolutist EO score and ART-G total score

| Predictor Variable | B(SE) | 95% CI for B | | β | R^2 | ΔR^2 |
|--------------------|-------|--------------|----|---------|-------|--------------|
| | | LL | UL | | | |

| | | | | | | |
|---------------|-------------|-------|------|----------|-----|-----|
| Step 1 | | | | | .01 | .01 |
| Constant | 2.76 (0.11) | 2.54 | 2.98 | <.001*** | | |
| Total reading | 0.00 (0.00) | -0.01 | 0.00 | .117 | | |

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

No significant relationship was shown.

Next, nonfiction and fiction ART-G scores were added in steps of a hierarchical regression with absolutism.

Table 4: Hierarchical regression results for absolutist EO and ART-G fiction and nonfiction score

| Predictor Variable | B(SE) | 95% CI for B | | β | R^2 | ΔR^2 |
|--------------------|--------------|--------------|-------|----------|-------|--------------|
| | | LL | UL | | | |
| Step 1 | | | | | .04 | .04*** |
| Constant | 2.39 (0.09) | 2.21 | 2.57 | <.001*** | | |
| Nonfiction reading | 0.03 (0.01) | 0.01 | 0.05 | <.001*** | | |
| Step 2 | | | | | .14 | .10*** |
| Constant | 2.73 (0.10) | 2.52 | 2.93 | <.001*** | | |
| Nonfiction reading | 0.06 (0.01) | 0.04 | 0.08 | <.001*** | | |
| Fiction reading | -0.02 (0.00) | -0.03 | -0.01 | <.001*** | | |

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

Nonfiction had a significant positive association with absolutism, and this remained significant when fiction was added to the regression. With nonfiction controlled for, fiction exposure has a significant negative association with absolutism. However, it should be noted that the increase in average absolutism per nonfiction author recognised was very small ($b = 0.06$), and likewise for the decrease per additional fiction author ($b = -0.02$).

Multiplism. As both educational level and gender were shown to have significant associations with average multiplism scores, these variables were included as controls in the first step of the hierarchical regression analyses. As the other gender categories contained extremely small participant numbers, only male and female gender participants were included in these analyses ($n = 302$). First ART-G total scores were assessed.

- Table 5: Hierarchical regression results for multiplist EO and ART-G total score

| Predictor Variable | B(SE) | 95% CI for B | | β | R^2 | ΔR^2 |
|--------------------|--------------|--------------|-------|----------|-------|--------------|
| | | LL | UL | | | |
| Step 1 | | | | | .07 | .07*** |
| Constant | 3.74 (0.20) | 3.35 | 4.12 | <.001*** | | |
| Male | -0.48 (0.12) | -0.72 | -0.23 | <.001*** | | |
| Educational level | -0.14 (0.05) | -0.25 | -0.03 | .011* | | |
| Step 2 | | | | | .13 | .06*** |
| Constant | 4.00 (0.20) | 3.62 | 4.39 | <.001*** | | |
| Male | -0.47 (0.12) | -0.71 | -0.24 | <.001*** | | |
| Educational level | -0.12 (0.05) | -0.22 | -0.01 | .026* | | |
| Total reading | -0.01 (0.00) | -0.02 | -0.01 | <.001*** | | |

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

Men had reduced average multiplism levels in contrast with women. Furthermore, as educational level increased, average multiplism levels also decreased. When ART-G total scores were added to the regression, model fit was significantly improved, and with both education and gender accounted for increased total ART-G score was associated with a very slight decrease in multiplism ($b = -0.01$).

Next, the hierarchical regression process was repeated with nonfiction and then fiction added in steps.

Table 6: Hierarchical regression results for multiplist EO and fiction and nonfiction score

| Predictor Variable | B(SE) | 95% CI for B | | β | R^2 | ΔR^2 |
|--------------------|-------|--------------|----|---------|-------|--------------|
| | | LL | UL | | | |
| Step 1 | | | | | .07 | .07*** |

| | | | | | | |
|--------------------|--------------|-------|-------|----------|-----|--------|
| Constant | 3.74 (0.20) | 3.35 | 4.12 | <.001*** | | |
| Male | -0.48 (0.12) | -0.72 | -0.23 | <.001*** | | |
| Educational level | 0.14 (0.05) | -0.25 | -0.03 | .011* | | |
| Step 2 | | | | | .22 | .15*** |
| Constant | 4.00 (0.18) | 3.65 | 4.36 | <.001*** | | |
| Male | -0.27 (0.12) | -0.50 | -0.04 | .021* | | |
| Educational level | -0.12 (0.05) | -0.22 | -0.02 | .017* | | |
| Nonfiction reading | -0.05 (0.01) | -0.07 | -0.04 | <.001*** | | |
| Step 3 | | | | | .21 | -.01 |
| Constant | 3.98 (0.19) | 3.61 | 4.35 | <.001*** | | |
| Male | -0.25 (0.12) | -0.49 | -0.02 | .035* | | |
| Educational level | -0.12 (0.05) | -0.22 | -0.02 | .015* | | |
| Nonfiction reading | -0.06 (0.01) | -0.07 | -0.04 | <.001*** | | |
| Fiction reading | 0.00 (0.00) | 0.00 | 0.01 | .538 | | |

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

When educational level and gender was controlled for, nonfiction ART-G score was significantly associated with lower average multiplism ($b = 0.06$). Adding fiction reading to the regression did not improve model fit, and fiction had no significant association with multiplism levels.

Evaluativism. As age and educational level were found to have significant associations with average evaluativism scores, these variables were included as controls in the first step of the hierarchical regression analyses. First, ART-G total scores were assessed.

Table 7: Hierarchical regression results for evaluativist EO and ART-G total score

| Predictor Variable | B(SE) | 95% CI for B | | β | R^2 | ΔR^2 |
|--------------------|-------------|--------------|------|----------|-------|--------------|
| | | LL | UL | | | |
| Step 1 | | | | | .04 | .04** |
| Constant | 4.62 (0.15) | 4.33 | 4.91 | <.001*** | | |
| Age | 0.00 (0.00) | 0.00 | 0.01 | .183 | | |
| Educational level | 0.10 (0.04) | 0.03 | 0.17 | .005** | | |
| Step 2 | | | | | .08 | .04*** |
| Constant | 4.58 (0.15) | 4.29 | 4.86 | <.001*** | | |

| | | | | |
|-------------------|-------------|-------|------|----------|
| Age | 0.00 (0.00) | -0.01 | 0.01 | .971 |
| Educational level | 0.10 (0.04) | 0.03 | 0.17 | .006** |
| Total reading | 0.01 (0.00) | 0.00 | 0.01 | <.001*** |

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

Educational level had a significant association with higher levels of evaluativism. When ART-G total scores were added to the regression, model fit was improved. When age and educational level were accounted for, the total ART-G score was associated with a very slight increase in average evaluativism ($b = 0.01$).

Next, nonfiction and fiction were added in steps to the hierarchical regression.

Table 8: Hierarchical regression results for evaluativist EO and fiction and nonfiction score

| Predictor Variable | B(SE) | 95% CI for B | | β | R^2 | ΔR^2 |
|--------------------|-------------|--------------|------|----------|-------|--------------|
| | | LL | UL | | | |
| Step 1 | | | | | .04 | .04** |
| Constant | 4.62 (0.15) | 4.33 | 4.91 | <.001*** | | |
| Age | 0.00 (0.00) | 0.00 | 0.01 | .183 | | |
| Educational level | 0.10 (0.04) | 0.03 | 0.17 | .005** | | |
| Step 2 | | | | | .04 | .00 |
| Constant | 4.61 (0.15) | 4.32 | 4.90 | <.001*** | | |
| Age | 0.00 (0.00) | 0.00 | 0.01 | .383 | | |
| Educational level | 0.10 (0.04) | 0.03 | 0.17 | .005** | | |
| Nonfiction reading | 0.01 (0.01) | 0.00 | 0.02 | .234 | | |
| Step 3 | | | | | .09 | .05*** |
| Constant | 4.57 (0.14) | 4.29 | 4.86 | <.001*** | | |
| Age | 0.00 (0.00) | -0.01 | 0.01 | .876 | | |
| Educational level | 0.09 (0.04) | 0.03 | 0.16 | .008** | | |
| Nonfiction reading | 0.00 (0.01) | -0.01 | 0.01 | .694 | | |
| Fiction reading | 0.01 (0.00) | 0.00 | 0.01 | <.001*** | | |

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

Adding nonfiction scored did not improve model fit, and there was no significant association between nonfiction and average evaluativism levels. When fiction was added, model fit was

significantly improved. Controlling for age, educational level, and nonfiction score, fiction ART-G scores were associated with a very slight increase in average evaluativism levels ($b = 0.01$).

Mediation by NT Upon EO. The significant relationships found between ART-G total, fiction and nonfiction scores and the different EO types were tested for mediation by NT. Model-based causal mediation analysis with nonparametric bootstrapped (1000 simulations) confidence level estimation using the percentile method was performed to calculate the average causal mediation effect.

Mediation Upon Absolutism. Results from the causal mediation analysis showed that nonfiction scores had an insignificant direct effect of 0.02 (95% CI 0.00, 0.2; $p = .124$), a significant indirect effect via NT of 0.01 (95% CI 0.00, 0.01; $p = .002$), and an insignificant total effect of 0.03 (95% CI -0.00, 0.05; $p = .09$). Therefore, NT was shown to fully mediate the relationship between nonfiction and absolutism.

Results from the causal mediation analysis showed that fiction scores had a significant direct effect of -0.01 (95% CI -0.02, 0.0; $p = .008$), a significant indirect effect via NT of -0.01 (95% CI -0.01, 0.0; $p = .004$), and a significant total effect of -0.01 (95% CI -0.02, 0.0; $p < .001$). The proportion of the total effect mediated by NT was 0.20 (95% CI 0.05, 0.5). Therefore, NT was shown to mediate the relationship between fiction and absolutism.

Mediation Upon Multiplism. Results from the causal mediation analysis showed that nonfiction scores had a significant direct effect of -0.07 (95% CI -0.07, -0.04; $p < .001$), an insignificant indirect effect via NT of -0.00 (95% CI -0.01, 0.00; $p = .42$), and a significant total effect of -0.06 (95% CI -0.06, -0.05; $p < .01$). Therefore, NT was shown not to mediate the relationship between nonfiction and multiplism.

Mediation Upon Evaluativism. Results from the causal mediation analysis showed that fiction scores had a significant direct effect of 0.01 (95% CI 0.00, 0.01; $p < .001$), a significant indirect effect via NT 0.00 (95% CI 0.00, 0.00; $p = .002$), and a significant total effect of 0.01 (95% CI 0.01, 0.01; $p = .002$). The proportion of the total effect mediated by NT was 0.14 (95% CI 0.03, 0.31). Therefore, NT was shown to mediate the relationship between fiction and evaluativism.

Discussion

This study investigated the impact of total, fiction and nonfiction print exposure on CT disposition and EO, taking account of NT. When educational level was accounted for, ART-G total scores did not predict increased CT disposition. When fiction and nonfiction were differentiated, opposite effects were found. Nonfiction exposure was significantly associated with lower CT disposition while fiction exposure was significantly associated with higher CT disposition. NT partly mediated the relationship between fiction and CT disposition, suggesting that reading more fiction in a highly transported state enhances disposition to think critically. The nonfiction relationship with CT disposition was fully mediated by NT, implying that the reduction in disposition towards CT arises from higher levels of transportation. These findings support the hypothesised relationship between fiction reading and CT disposition. The hypothesised relationship between total reading and CT disposition was not supported. However, in the case of nonfiction, an increased CT disposition was also hypothesised, albeit with a smaller effect than fiction; the findings show that nonfiction and fiction in fact have opposite directions of effect.

The findings pertaining to absolutist EO showed that ART-G total scores had no significant relationship to average absolutism levels. Nonfiction exposure had a significant positive relationship

with absolutism. This relationship was fully mediated by NT, implying the increased absolutism results from high levels of transportation in nonfiction reading. Fiction exposure had the inverse relationship, with higher levels of fiction significantly associated with lower absolutism. This was partly mediated by NT, suggesting highly transported fiction reading enhances the reduction in absolutism. Addressing multiplist EO, gender and educational level were controlled for, and total ART-G score was found to have a significant negative relationship with multiplism. Separating out fiction and nonfiction, only nonfiction was found to have a significant negative association. Neither of these relationships was mediated by NT. Finally, addressing evaluativist EO, controlling for age and educational level, ART-G total scores had a significant positive relationship with evaluativism, with no mediation by NT. When separating out fiction and nonfiction, only fiction exposure had a significant positive association with evaluativism. This relationship was partly mediated by NT, suggesting a heightened effect through greater transportation.

In framing the hypotheses pertaining to EO, the developmental trajectory stipulated in the literature from absolutism, through multiplism and into evaluativism, was taken as a move from less to more sophisticated EO (Hofer & Pintrich, 1997), with evaluativists in the most sophisticated EO category. As evaluativist EO is also considered necessary for CT (Kuhn et al., 2000), this framing fitted the aims of the study in using EO as an indicator of CT. As such, it was hypothesised that increased reading would be associated with increased evaluativism, especially when it came to fiction. Higher total, and fiction, scores were associated with increased average evaluativism; these findings support the hypotheses. However, for nonfiction, the hypothesis is not supported as nonfiction scores had no significant relationship with evaluativism. Addressing the other two EO types, the findings are more complex. In terms of multiplism, both total and nonfiction exposure were associated with lower levels, in keeping with the hypotheses. However, fiction exposure had no association with multiplism. As it was hypothesised that fiction would have a stronger effect on EO than nonfiction, this hypothesis is not supported in terms of multiplism. Finally, when it came to absolutism only fiction exposure had a significant relationship with lower average absolutism. This

supports the hypotheses to the extent that fiction was found to have a stronger effect but not in the sense that all reading was hypothesised to have some effect, upon EO.

Finally, it was hypothesised that NT would mediate the effects of reading on CT disposition and EO. This hypothesis is supported in some cases as detailed above, but not all.

Wider Issues and Limitations

ART-G scores give an indication of the extent to which participants are exposed to author names but do not measure actual reading levels, and thus the results of this study do not capture actual reading behaviour. Furthermore, the ART-G was designed and validated in Canada, and therefore the authors used in it may be more familiar to a Canadian audience than participants from other countries. Participants for this study were partly recruited online, and thus could be English speakers from any part of the world; however, the limitation of the prize voucher to the UK likely also limited the sample. Furthermore, the study was advertised with posters in London public spaces such as community centres, cafes, and libraries, again increasing the likelihood of a UK-based sample. While the names on the ART-G list are international, not focused on Canadian authors, there are differences in the UK and Canadian book markets; for example, comparing the top 10 bestselling books of 2019, only one book *The Tattooist of Auschwitz* featured in both countries' lists (Dugdale, 2019; Sparkles, 2019). Such bookseller market differences could result in differences in the familiarity of some authors between the two countries. It should be noted, however, that the potential difference in familiarity of authors between the two countries ought not disadvantage some participants against others in this study, as all are likely to be UK based, and therefore the comparative scores between participants within this sample can still be considered representative of their exposure to fiction and nonfiction. Still, this may be one explanation for distributions of scores skewing low (figures 1-3). Additionally, the ART-G contains many more fiction than nonfiction authors, and though it permits division into subscales for fiction genres, it does not have subscales for nonfiction topics. Thus the measure is arguably more sensitive for fiction than nonfiction.

Further, to the limitations of the ART-G, differentiating fiction and nonfiction is not a straightforward matter. These are not mutually exclusive categories that can be given necessary and sufficient conditions to clearly delineate them, and many of the devices associated with fiction are often deployed in nonfiction writing (Friend, 2008). Indeed, it has been argued that both nonfiction and fiction can be classed as 'literature' (Searle, 1975), or further that there is a distinct class of 'narrative nonfiction,' 'literary nonfiction' or even the 'nonfiction novel' (Friend, 2008; Heyne, 1987). These categories point to the use of narrative in nonfiction writing, which can utilise colourful storytelling in depicting nonfictional events. As Friend (2008) points out, the creative process of writing can be very much the same in fiction and nonfiction, and the reader's experience can be as rich in either. Conversely, it is not the case that only nonfiction texts convey non-fictional information, as fictional works can also contain elements that are factual and not presented in a narrative; for example, the novel *Seveneves* by Neal Stephenson contains long passages relating information about orbital mechanics with a high level of thoroughness and accuracy (Elvik, 2018). This problematises the division of fiction and nonfiction with expected differences in effects, as both may build knowledge in some cases, and both may be highly transporting into narratives in some instances. It is noteworthy that the effects of nonfiction were fully mediated by NT for CT disposition, and for absolutist EO. This suggests that increased nonfiction reading was connected to lower disposition to think critically and higher absolutism only with high levels of transportation. Indeed, other researchers attributed differences in fiction and nonfiction outcomes to their narrative and non-narrative nature (e.g., Mar et al., 2006), although authors included in the ART nonfiction list capture both (e.g., Oliver Sacks is a nonfiction author whose books contain accounts of patient experiences which are often highly narrative in style). It would be interesting therefore to test whether different kinds of nonfiction materials(i.e., those that are more or less narrative driven) would have different impacts. This study presents some broad correlations that differentiate fiction and nonfiction in a generalised manner; future work would benefit from more in-depth focus on different forms of fiction and nonfiction material.

Further, to the limitation in separating out fiction and nonfiction, with different levels of narrativity present in different nonfiction texts, and some non-narrative elements in fiction texts, there are other differences in text types and purposes that are likely relevant to their relationship with CT. As was discussed in the introduction, NT is often associated with reduced CT; however, this research typically focuses on narratives designed for persuasion, such as advertising. In this domain, being highly transported by a marketing narrative reduces CT while increasing affect, leading to more emotional and less critical brand evaluations (Escalas, 2004). It could be argued that the reduced disposition to think critically associated with nonfiction and its full mediation by NT found in this study captures the same effect. However, the ART-G features authors of nonfiction books, not advertising. Some of these authors may indeed be known for persuasive writing leveraging narrative to prompt reduced criticality, but many in the scale, such as Stephen Hawking or Susan Sontag for example, are known for works it would be torturous to compare with advertising. Similarly, the findings pertaining to fiction that were mediated by NT in this study cannot be said to contradict those linking NT to reduced CT, due the difference in the nature of the reading material. Therefore, while the prior research into NT and critical thought offers some broader context to these findings, there is a mismatch in the nature of the reading materials in consideration, and therefore conclusions cannot be drawn across these domains. Further research differentiating different types of materials authored for different purposes is needed to explore the relationship between reading, NT and CT in detail.

EO is typically conceptualised in developmental terms as a progression from absolutism, through multiplism, to evaluativism (Hofer & Pintrich, 1997). However, given variation in adult EO (Kuhn et al., 2000), it may be better thought of as distinct but not necessarily ordered categories. Furthermore, this developmental ranking may be biased towards one way of ordering these orientations from a Western cultural perspective. Cross-cultural studies comparing epistemological orientation are somewhat rare, with some finding differences (Tabak & Weinstock, 2008), while others do not (Zeidler et al., 2013). Nonetheless, a classification of these orientations as better or

worse than one another seems unlikely to have universalisability. For the purposes of this research, evaluativism has been identified as most conducive to CT, but this does not mean that absolutism- or multiplism-oriented individuals cannot think critically. Therefore, conclusions about different reading associations with different levels of EO should be taken as indicative of different underpinning stances that could inform CT. This study shows that reading fiction and nonfiction is associated with different positions in the EO space. Further qualitative research would be valuable in providing more detail and nuance to how different reading habits may inform different orientations towards epistemology and how these then can relate to different CT approaches.

The outcome variables in this study were measured using self-report Likert scales and, as such, must be interpreted as subjective to participants' perceptions of themselves. It is likely that desirability effects played a role in responses to both scales. The CTDS has items such as "I am often on the lookout for new ideas" (Sosu, 2013), which carry a clearly positive connotation. The JCI is perhaps less prone to desirability effect as the statements, such as "Most debatable issues have one right conclusion" (McGinnis, 2016), are less normative. However, both scales are validated, and participants in the sample had diverse responses to both. Therefore, although these reflective subjective self-assessments, they are still useful indicators of differences between participants. Additionally, the scales were administered in the same sequence to all participants, and thus order effects may be present; as the JCI was presented last in the survey (and indeed some participants did not complete it), it may be that JCI responses were given with greater fatigue than the prior scales. However, as the JCI contains three subscales, it is arguably more varied and thus perhaps more interesting than the CTDS. For future studies, alternating scale ordering would be recommended to avoid this potential issue.

It is not possible to determine a causal direction based upon these findings. It may be that people with higher CT disposition chose to read more fiction, rather than that reading improves CT disposition. Similarly, people with more evaluativist rather than absolutist stances may be more drawn to fiction reading while people who are more absolutist may be more attracted to nonfiction.

The findings from this study merely demonstrate a relationship between these factors, but the direction of causal influence requires further investigation.

Conclusions

The main aim of this study was to identify whether there is a relationship between reading engagement and factors associated with CT and whether fiction and nonfiction had distinct associations. By addressing CT disposition and EO, an indication can be given of a further relationship between reading and CT as a broader construct, to be pursued in further research. While reading engagement measured through print exposure was not associated with differences in CT disposition, but was associated with different EO, fiction and nonfiction taken separately had distinctive relationships sometimes in opposite directions. Increased fiction exposure was associated with increased CT disposition, suggesting fiction readers are more disposed towards CT, while a negative association was found for nonfiction exposure. Fiction and nonfiction exposure were associated with different EO levels, suggesting that readers of these different types of materials have different epistemological beliefs that may feed into their CT approaches. Therefore, the findings of this study imply that reading fiction and nonfiction have different relationships with factors associated with CT, and therefore research into reading and CT would be a worthwhile endeavour.

References

- Abrami, P. C., Bernard, R. M., Borokhovski, E., Wade, A., Surkes, M. A., Tamim, R., & Zhang, D. (2008). Instructional Interventions Affecting Critical Thinking Skills and Dispositions: A Stage 1 Meta-Analysis. *Review of Educational Research, 78*(4), 1102–1134. <https://doi.org/10.3102/0034654308326084>
- Aloqaili, A. S. (2012). The relationship between reading comprehension and critical thinking: A theoretical study. *Journal of King Saud University - Languages and Translation, 24*(1), 35–41. <https://doi.org/10.1016/j.jksult.2011.01.001>
- Appel, M., & Richter, T. (2007). Persuasive effects of fictional narratives increase over time. *Media Psychology, 10*(1), 113–134. <https://doi.org/10.1080/15213260701301194>
- Arslan, S., & Demirtas, Z. (2016). Social emotional learning and critical thinking disposition. *Studia Psychologica, 58*(4), 276–285. <https://doi.org/10.21909/sp.2016.04.723>
- Avin, S. (2019). Exploring artificial intelligence futures. *Journal of AI Humanities, 2*, 171–193. <https://doi.org/10.17863/CAM.35812>
- Bal, M. P., & Veltkamp, M. (2013). How does fiction reading influence empathy? An experimental investigation on the role of emotional transportation. *PLoS ONE, 8*(1), e55341. <https://doi.org/10.1371/journal.pone.0055341>
- Batchelor, O. (2017). Getting out the truth: The role of libraries in the fight against fake news. *Reference Services Review, 45*(2), 143–148. <https://doi.org/10.1108/RSR-03-2017-0006>
- Black, J. E., Capps, S. C., & Barnes, J. L. (2017). Fiction, genre exposure, and moral reality. *Psychology of Aesthetics, Creativity, and the Arts*. <https://doi.org/10.1037/aca0000116>
- Blanca, M. J., Alarcón, R., & Arnau, J. (2017). Non-normal data: Is ANOVA still a valid option? *Psicothema, 29.4*, 552–557. <https://doi.org/10.7334/psicothema2016.383>
- Busselle, R., & Bilandzic, H. (2008). Fictionality and perceived realism in experiencing stories: A Model of narrative comprehension and engagement. *Communication Theory, 18*(2), 255–280. <https://doi.org/10.1111/j.1468-2885.2008.00322.x>

- Byrne, R. M. J. (2016). Counterfactual thought. *Annual Review of Psychology*, *67*(1), 135–157.
<https://doi.org/10.1146/annurev-psych-122414-033249>
- Byrnes, J. P., & Dunbar, K. N. (2014). The nature and development of critical-analytic thinking. *Educational Psychology Review*, *26*(4), 477–493. <https://doi.org/10.1007/s10648-014-9284-0>
- Cacioppo, J. T., & Petty, R. E. (1984). The elaboration likelihood model of persuasion. In T. E. Kinnear (Ed.), *Advances in Consumer Research* (Vol. 11, pp. 673–675). Association for Consumer Research.
- Chan, C. (2019). Using digital storytelling to facilitate critical thinking disposition in youth civic engagement: A randomized control trial. *Children and Youth Services Review*, *107*, 104522.
<https://doi.org/10.1016/j.childyouth.2019.104522>
- Colucciello, M. L. (1997). Critical thinking skills and dispositions of baccalaureate nursing students—A conceptual model for evaluation. *Journal of Professional Nursing*, *13*(4), 236–245.
[https://doi.org/10.1016/S8755-7223\(97\)80094-4](https://doi.org/10.1016/S8755-7223(97)80094-4)
- Corballis, M. C. (2009). Mental time travel and the shaping of language. *Experimental Brain Research*, *192*(3), 553–560. <https://doi.org/10.1007/s00221-008-1491-9>
- Cuff, B. M. P., Brown, S. J., Taylor, L., & Howat, D. J. (2014). Empathy: A Review of the Concept. *Emotion Review*, *8*(2), 144–153. <https://doi.org/10.1177/1754073914558466>
- Dal Cin, S., Zanna, M. P., & Fong, G. T. (2004). Narrative persuasion and overcoming resistance. In E. S. Knowles & J. A. Linn (Eds.), *Resistance and persuasion*. (pp. 175–191). Lawrence Erlbaum Associates.
- Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, *44*(1), 113–126.
<https://doi.org/10.1037/0022-3514.44.1.113>
- DioGuardi, S. (2016). Critical thinking in criminal justice ethics: Using the affective domain to discover gray matters. *Journal of Criminal Justice Education*, *27*(4), 535–548.
<https://doi.org/10.1080/10511253.2016.1145709>

- Djikic, M., Oatley, K., & Moldoveanu, M. C. (2013). Opening the closed mind: The effect of exposure to literature on the need for closure. *Creativity Research Journal*, 25(2), 149–154.
<https://doi.org/10.1080/10400419.2013.783735>
- Djikic, M., Oatley, K., Zoeterman, S., & Peterson, J. B. (2009). Defenseless against art? Impact of reading fiction on emotion in avoidantly attached individuals. *Journal of Research in Personality*, 43(1), 14–17. <https://doi.org/10.1016/j.jrp.2008.09.003>
- Dugdale, J. (2019, December 31). Bestselling books of 2019. *The Guardian*.
<https://www.theguardian.com/books/2019/dec/31/bestselling-books-2019>
- Dumitru, D. (2019). Creating meaning: The importance of arts, humanities and culture for critical thinking development. *Studies in Higher Education*, 44(5), 870–879.
<https://doi.org/10.1080/03075079.2019.1586345>
- Ekinci, Ö., & Aybek, B. (2010). Analysis of the empathy and the critical thinking disposition of the teacher candidates. *Ilkogretim Online*, 9(2), 816–816.
- Elvik, S. (2018). The orbits of Seveneves: A book review with a touch of orbital dynamics. *Principium*, 20(February), 16–18.
- Ennis, R. H. (2016). Definition: A three-dimensional analysis with bearing on key concepts. *OSSA Conference Archive*, 115.
https://scholar.uwindsor.ca/ossaarchive/OSSA11/papersandcommentaries/105?utm_source=scholar.uwindsor.ca%2Fossaarchive%2FOSSA11%2Fpapersandcommentaries%2F105&utm_medium=PDF&utm_campaign=PDFCoverPages
- Escalas, J. E. (2004). Imagine yourself in the product: Mental simulation, narrative transportation, and persuasion. *Journal of Advertising*, 33(2), 37–48.
<https://doi.org/10.1080/00913367.2004.10639163>

- Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction. Research findings and recommendations.*
<https://eric.ed.gov/?id=ED315423>
- Fong, K., Mullin, J. B., & Mar, R. A. (2015). How exposure to literary genres relates to attitudes toward gender roles and sexual behavior. *Psychology of Aesthetics, Creativity, and the Arts, 9*, 274–285. <https://doi.org/10.1037/a0038864>
- Friend, S. (2008). Imagining fact and fiction. In K. Stock & K. Thomson-Jones (Eds.), *New waves in aesthetics* (pp. 150–169). Palgrave Macmillan.
- Gerrig, R. J. (1993). *Experiencing narrative worlds: On the psychological activities of reading.* Yale University Press. <https://doi.org/10.2307/j.ctt1dr37cb>
- Greco, J. (2000). Two kinds of intellectual virtue. *Philosophy and Phenomenological Research, 60*(1), 179–184. <https://doi.org/10.2307/2653438>
- Green, M. C., & Brock, T. C. (2013). In the mind's eye: Transportation-imagery model of narrative persuasion. In M. C. Green, J. J. Strange, & T. C. Brock (Eds.), *Narrative impact: Social and cognitive foundations* (pp. 287–313). Psychology Press.
- Hakemulder, J. (2001). How to make alle menschen bruder: Literature in a multicultural and multiform society. In G. Steen, D. H. Schram, & E. Ibsch (Eds.), *The psychology and sociology of literature: In honor of Elrud Ibsch* (pp. 225–242). John Benjamins.
- Hakemulder, J. (2008). Imagining what could happen: Effects of taking on the role of a character on social cognition. In S. Zyngier, M. Bortolussi, A. Chesnokova, & J. Auracher (Eds.), *Directions in empirical literary studies: In honor of Willie van Peer* (pp. 139–153). John Benjamins.
- Hamby, A., Brinberg, D., & Daniloski, K. (2017). Reflecting on the journey: Mechanisms in narrative persuasion. *Journal of Consumer Psychology, 27*(1), 11–22.
<https://doi.org/10.1016/j.jcps.2016.06.005>

- Hampson Lundh, A., Dolatkah, M., & Limberg, L. (2018). From informational reading to information literacy. *Journal of Documentation*, 74(5), 1042–1052. <https://doi.org/10.1108/JD-11-2017-0156>
- Hawkins, K. (2012). *Thinking and reading among college undergraduates: An examination of the relationship between critical thinking skills and voluntary reading* [Doctoral dissertation, University of Tennessee]. https://trace.tennessee.edu/utk_graddiss/1302
- Heyne, E. (1987). Toward a theory of literary nonfiction. *MFS Modern Fiction Studies*, 33(3), 479–490. <https://doi.org/10.1353/mfs.0.1150>
- Hoeken, H., & Flikkers, K. M. (2014). Issue-relevant thinking and identification as mechanisms of narrative persuasion. *Poetics*, 44, 84–99. <https://doi.org/10.1016/j.poetic.2014.05.001>
- Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, 67(1), 88–140. <https://doi.org/10.3102/00346543067001088>
- Holma, K. (2015). The critical spirit: Emotional and moral dimensions of critical thinking. *Studier i Pædagogisk Filosofi*, 4(1), 17–28. <https://doi.org/10.7146/spf.v4i1.18280>
- hooks, b. (1994). *Teaching to transgress: Education as the practice of freedom*. Routledge.
- Hyytinen, H., Holma, K., Toom, A., Shavelson, R. J., & Lindblom-Ylänne, S. (2014). The complex relationship between students' critical thinking and epistemological beliefs in the context of problem solving. *Frontline Learning Research*, 2(5), 1–25. <https://doi.org/10.14786/flr.v2i4.124>
- Jeong, H. S. (2015). Critical thinking disposition, problem solving process, and empathy among nursing students. *Advanced Science and Technology Letters*, 103, 44–48. <https://doi.org/10.14257/astl.2015.103.10>
- Johnson, D. R. (2012). Transportation into a story increases empathy, prosocial behavior, and perceptual bias toward fearful expressions. *Personality and Individual Differences*, 52(2), 150–155. <https://doi.org/10.1016/j.paid.2011.10.005>

- Kidd, D. C., & Castano, E. (2013). Reading literary fiction improves theory of mind. *Science*, 342(6156), 377–380. <https://doi.org/10.1126/science.1239918>
- King, P. M., & Kitchener, K. S. (2004). Reflective judgment: Theory and research on the development of epistemic assumptions through adulthood. *Educational Psychologist*, 39(1), 5–18. https://doi.org/10.1207/s15326985ep3901_2
- Koek, M., Janssen, T., Hakemulder, F., & Rijlaarsdam, G. (2016). Literary reading and critical thinking: Measuring students' critical literary understanding in secondary education. *Scientific Study of Literature*, 6(2), 243–277. <https://doi.org/10.1075/ssol.6.2.04koe>
- Kuhn, D. (1991). *The skills of argument*. Cambridge University Press.
- Kuhn, D. (2019). Critical thinking as discourse. *Human Development*, 62(3), 146–164. <https://doi.org/10.1159/000500171>
- Kuhn, D., Cheney, R., & Weinstock, M. (2000). The development of epistemological understanding. *Cognitive Development*, 15(3), 309–328. [https://doi.org/10.1016/S0885-2014\(00\)00030-7](https://doi.org/10.1016/S0885-2014(00)00030-7)
- Liu, S.-Y., Lin, C.-S., & Tsai, C.-C. (2011). College students' scientific epistemological views and thinking patterns in socioscientific decision making. *Science Education*, 95(3), 497–517. <https://doi.org/10.1002/sce.20422>
- Lutzke, L., Drummond, C., Slovic, P., & Árvai, J. (2019). Priming critical thinking: Simple interventions limit the influence of fake news about climate change on Facebook. *Global Environmental Change*, 58, 101964. <https://doi.org/10.1016/j.gloenvcha.2019.101964>
- Mar, R. A., Oatley, K., Hirsh, J., dela Paz, J., & Peterson, J. B. (2006). Bookworms versus nerds: Exposure to fiction versus non-fiction, divergent associations with social ability, and the simulation of fictional social worlds. *Journal of Research in Personality*, 40, 694–712. <https://doi.org/10.1016/j.jrp.2005.08.002>
- Mar, R. A., Oatley, K., & Peterson, J. B. (2009). Exploring the link between reading fiction and empathy: Ruling out individual differences and examining outcomes. *Communications*, 34(4), 407–428. <https://doi.org/10.1515/COMM.2009.025>

- Mar, R. A., & Rain, M. (2015). Narrative fiction and expository nonfiction differentially predict verbal ability. *Scientific Studies of Reading, 19*(6), 1–15.
<https://doi.org/10.1080/10888438.2015.1069296>
- Mason, L., Krutka, D., & Stoddard, J. (2018). Media literacy, democracy, and the challenge of fake news. *Journal of Media Literacy Education, 10*(2), 1–10. <https://doi.org/10.23860/JMLE-2018-10-2-1>
- Mazzocco, P. J., Green, M. C., Sasota, J. A., & Jones, N. W. (2010). This story is not for everyone: Transportability and narrative persuasion. *Social Psychological and Personality Science, 1*(4), 361–368. <https://doi.org/10.1177/1948550610376600>
- McGinnis, D. (2016). Epistemological orientations and evidence evaluation in undergraduates. *Thinking Skills and Creativity, 19*, 279–289. <https://doi.org/10.1016/j.tsc.2016.01.002>
- Menninghaus, W., Wagner, V., Hanich, J., Wassiliwizky, E., Jacobsen, T., & Koelsch, S. (2017). The Distancing-Embracing model of the enjoyment of negative emotions in art reception. *Behavioral and Brain Sciences, 40*, e347. <https://doi.org/10.1017/S0140525X17000309>
- Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *The Behavioral and Brain Sciences, 34*(2), 57–74.
<https://doi.org/10.1017/S0140525X10000968>
- 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.
<https://doi.org/10.1037/aca0000089>
- Oatley, K. (2011). *Such stuff as dreams: The psychology of fiction*. Wiley-Blackwell.
- Özyürek, A., & Trabasso, T. (1997). Evaluation during the understanding of narratives. *Discourse Processes, 23*(3), 305–335. <https://doi.org/10.1080/01638539709544995>
- Pardede, P. (2019). Using fiction to promote students' critical thinking. *JET (Journal of English Teaching), 5*(3), 166–178. <https://doi.org/10.33541/jet.v5i3.1309>

- Paul, R. (1987). Dialogical thinking: Critical thought essential to the acquisition of rational knowledge and passions. In J. B. Baron & R. J. Sternberg (Eds.), *Teaching thinking skills: Theory and practice* (pp. 127–148). W. H. Freeman and Company.
- Pu, D., Ni, J., Song, D., Zhang, W., Wang, Y., Wu, L., Wang, X., & Wang, Y. (2019). Influence of critical thinking disposition on the learning efficiency of problem-based learning in undergraduate medical students. *BMC Medical Education, 19*(1), 1. <https://doi.org/10.1186/s12909-018-1418-5>
- Richmond, S. (2004). Being in others: Empathy from a psychoanalytical perspective. *European Journal of Philosophy, 12*(2), 244–264. <https://doi-org/10.1111/j.0966-8373.2004.00209.x>
- Searle, J. R. (1975). The logical status of fictional discourse. *New Literary History, 6*(2), 319–332. <https://doi.org/10.2307/468422>
- Serpell, C. N. (2010). Repetition and the ethics of suspended reading in American Psycho. *Critique: Studies in Contemporary Fiction, 51*(1), 47–73. <https://doi.org/10.1080/00111610903249864>
- Shihab, I. A. (2011). Reading as critical thinking. *Asian Social Science, 7*(8), 209–218. <https://doi.org/10.5539/ass.v7n8p209>
- Sosu, E. M. (2013). The development and psychometric validation of a Critical Thinking Disposition Scale. *Thinking Skills and Creativity, 9*, 107–119. <https://doi.org/10.1016/j.tsc.2012.09.002>
- Sparkles, A. (2019, December 9). Bestselling books of 2019. *Booknet Canada*. <https://www.booknetcanada.ca/blog/2019/12/9/2019-best-sellers>
- Stansfield, J., & Bunce, L. (2014). The relationship between empathy and reading fiction: Separate roles for cognitive and affective components. *Journal of European Psychology Students, 5*, 9–18.
- Strømsø, H. I., Bråten, I., & Samuelstuen, M. S. (2008). Dimensions of topic-specific epistemological beliefs as predictors of multiple text understanding. *Learning and Instruction, 18*(6), 513–527. <https://doi.org/10.1016/j.learninstruc.2007.11.001>

- Tabačková, Z. (2015). Outside the classroom thinking inside the classroom walls: Enhancing students' critical thinking through reading literary texts. *The Proceedings of 5th World Conference on Learning, Teaching and Educational Leadership*, 186, 726–731.
<https://doi.org/10.1016/j.sbspro.2015.04.042>
- Tabak, I., & Weinstock, M. (2008). A sociocultural exploration of epistemological beliefs. In M. S. Khine (Ed.), *Knowing, Knowledge and Beliefs: Epistemological Studies across Diverse Cultures* (pp. 177–195). Springer Netherlands. https://doi.org/10.1007/978-1-4020-6596-5_8
- Tamir, D. I., Bricker, A. B., Dodell-Feder, D., & Mitchell, J. P. (2016). Reading fiction and reading minds: The role of simulation in the default network. *Social Cognitive and Affective Neuroscience*, 11(2), 215–224. <https://doi.org/10.1093/scan/nsv114>
- Thayer-Bacon, B. J. (1992). A feminine reconceptualization of critical thinking theory. *Journal of Thought*, 27(1/2), 4–17.
- Thayer-Bacon, B. J. (2000). *Transforming critical thinking: Thinking constructively*. Teachers College Press.
- Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2014). mediation: R package for causal mediation analysis. *Journal of Statistical Software*, 59(5).
<https://doi.org/10.18637/jss.v059.i05>
- Tsui, L. (2000). Effects of campus culture on students' critical thinking. *The Review of Higher Education*, 23(4), 421–441. <https://doi.org/doi:10.1353/rhe.2000.0020>
- van Laer, T., de Ruyter, K., Visconti, L. M., & Wetzels, M. (2013). The extended transportation-imagery model: A meta-analysis of the antecedents and consequences of consumers' narrative transportation. *Journal of Consumer Research*, 40(5), 797–817.
<https://doi.org/10.1086/673383>
- Weinstock, M., Kienhues, D., Feucht, F. C., & Ryan, M. (2017). Informed reflexivity: Enacting epistemic virtue. *Educational Psychologist*, 52(4), 284–298.
<https://doi.org/10.1080/00461520.2017.1349662>

Zunshine, L. (2006). *Why we read fiction: Theory of mind and the novel*. Ohio State University Press.