UNIVERSITY OF BIRMINGHAM

‘What are They Feeling Now?’ The role of Theory of Mind but not Inhibitory Control in Individual Differences of Adolescent Emotional Inferencing Abilities

[redacted]

Word count: 4,413/4500
(granted 500 additional words)

Submitted in partial fulfilment

of the requirements for the degree of

Bachelor of Science

School of Psychology

University of Birmingham

Abstract

Emotional inferencing, a component of reading comprehension, is an important skill in understanding narrative literature. To infer a character's emotional state, readers form mental models containing representations of the text and use these to make inferences. In this study, we explored the potential role of two cognitive factors in forming these mental models: theory of mind (ToM) and inhibitory control (IC). These abilities have been linked to language abilities, but their relationship with affective inferencing has yet to be studied. Furthermore, considering the importance of ToM in adolescence, and because cognitive abilities have been argued to be more important for reading comprehension in older students, we specifically investigated adolescent emotional inferencing abilities. We hypothesised that adolescents with better ToM and IC will form better mental models compared to those with poorer abilities. This would be reflected in their faster reading times when reading sentences congruent with an emotion established by context and longer reading times for incongruent sentences. This relationship will be independent of the effects of linguistic factors.

Two hundred and ninety-six 16- to 18-year-old students from United Kingdom sixth forms completed questionnaires measuring the cognitive and linguistic factors, followed by a self-paced reading task. Results suggested that ToM could be involved in updating mental models rather than forming them. On the other hand, IC was not found to be predictive of emotional-inferencing abilities. However, we propose that this is due to how we conceptualised and measured IC. If, in replication studies, ToM and IC are found to be significant and relevant predictors of updating and forming mental models, these cognitive factors could be included in reading intervention research involving narrative texts with shifting emotions and perspectives. Finally, we discuss limitations to our study and recommend methodological changes for future emotional-inferencing research.

‘What are They Feeling Now?’ The role of Theory of Mind but not Inhibitory Control in Individual Differences of Adolescent Emotional Inferencing Abilities

Reading comprehension is the process of extracting and assembling meaning from text (Snow & Sweet, 2003). Unsurprisingly, it is an indispensable ability. For students, comprehension-test scores correlate with academic achievement (Savolainen et al., 2008) and poor reading has been linked to increased school dropout rates (Daniel et al., 2006). Thus, improving reading comprehension remains a critical interest in education.

One process contributing to text comprehension is inference generation (Kendeou et al., 2008). McKoon and Ratcliff (1992) defined inferencing as the interaction between information from the text and knowledge the reader already knows, producing a representation of the text, then using this to make deductions. Inferencing is a strong predictor of narrative comprehension in students (Oakhill & Cain, 2012) and Hall et al. (2020) found that improving students’ inferencing abilities increased their reading comprehension scores. Given its significance, our study will concentrate on inference abilities and factors affecting them, focusing on one specific type: emotional inferencing. This is the ability to deduce the affective state of a character in a narrative (Gillioz et al., 2012).

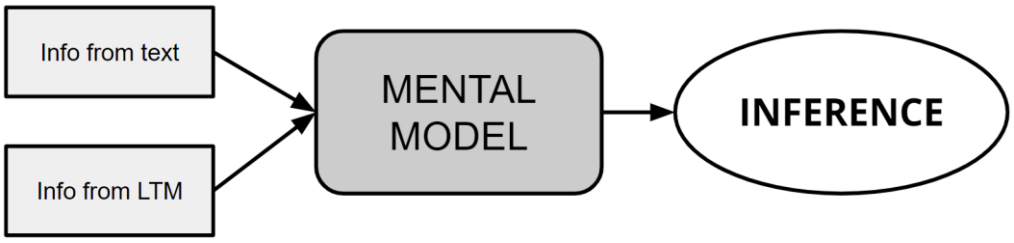
Emotional Inferences in Reading Narratives

Emotions are a vital part of everyday social interaction and discourse. The ability to understand the emotion implied by a situation is advantageous in understanding and inferring people’s behaviours (Gygax et al., 2003). The importance of emotional inferencing for reading comprehension is corroborated by the proposed mechanism behind inference generation: that readers construct mental situation models (Johnson-Laird, 1983; Van Dijk & Kintsch, 1983). When reading about a situation, readers incorporate information from the text

and their long-term memory (Gernsbacher, 1997) to create a model (see *Figure 1*). Using this model, readers can infer information from the narrative; in emotional inferencing, this is the affective state of the character, even when only implicitly implied. Take, for example: '*Mark had his job interview at the law firm last week. He trembled as he entered the interview room*.' Here, we infer that Mark is feeling *anxious*, even though this is never explicitly mentioned. Therefore, the ability to form a mental model and deduce this emotional state from context makes inferencing important in obtaining global coherence (the overall understanding) of an affective text (Graesser et al., 1994; Kneepkens & Zwaan, 1995).

Figure 1.

The Mental Model Mechanism of Inference Generation



Note. LTM = long-term memory

Gernsbacher et al. (1992) illustrated that readers form representations of a character's emotion using a self-paced reading paradigm. The authors showed that sentences containing emotions congruent with the context are read faster than sentences with incongruent emotions; this is the congruence effect. Upon reading the context, '*Mark had his job interview at the law firm last week,'* readers form the model that Mark is anxious. Readers then have no difficulty reading a congruent-emotion sentence ('*He trembled as he entered the interview room.'*) as this fits their model. However, reading times increase when reading an incongruent sentence ('*He chuckled as he entered the interview room.'*), reflecting the

reader's surprise as their model is violated. Regardless of whether the sentence contains positively-valenced (e.g. happy) or negatively-valenced (e.g. sad) emotions, incongruent sentences are read slower than congruent ones (Gygax et al., 2003). Thus, a reader's ability to form mental models is reflected in reading times of congruent versus incongruent sentences.

Individual Differences in Emotional Inferencing

The question remains however: what factors impact one's ability to form affective representations of narratives? Our study will investigate two potential cognitive abilities: theory of mind (ToM) and inhibitory control (IC). Further, three linguistic factors, print exposure, vocabulary knowledge and reading enjoyment, will be controlled to fully investigate the effects of the two abilities.

Theory of mind

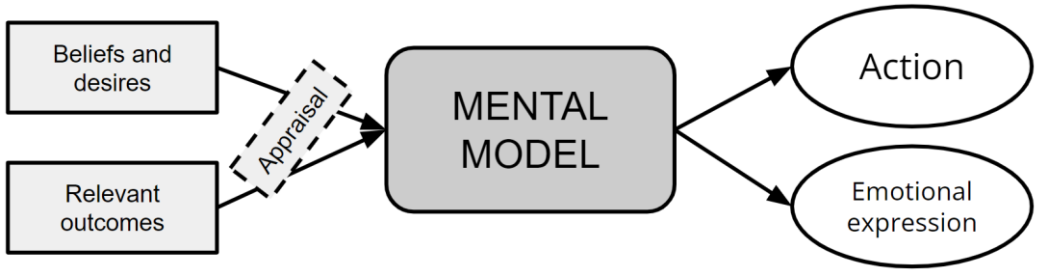
ToM, or 'mentalising', describes the ability to attribute mental states such as beliefs, desires and emotions to yourself and others (Premack & Woodruff, 1978). Successful implementation of this ability allows one to understand others' perspectives, predict their actions and understand their emotions (Luke & Banerjee, 2013; Xavier et al., 2013).

In their model, Ong et al. (2018) explored this link between emotion recognition and ToM (see *Figure 2*). They proposed that, in a situation, we (observers) first identify the agent's beliefs and desires, along with the situational outcomes relevant to them. Next, we use this information to 'appraise' the emotion they feel. Finally, we predict the emotional expression the agent will display. For Mark, we know from context and our knowledge that he wants the job and thus wants to do well in the interview. Therefore, we evaluate that he is feeling anxious. We can then predict that he will display actions reflecting this emotion and understand why he *trembles* as he enters the interview room. We can see an overlap between

this model and the mental model of inferencing; we use situational context and our long-term memory knowledge to appraise emotions and predict affective outcomes. Here, 'mentalising' is a key factor in inferring an agent's emotion.

Figure 2

Ong et al.'s (2018) Proposed Model of Emotional Inference in Theory of Mind



Research has supported this notion. Cutting and Dunn (1999) found significant correlations between individual differences in ToM and emotional understanding in children. Grazzani et al. (2018) confirm this, further reporting that emotional understanding explained variability in ToM scores. By teaching emotion recognition to children with autism, Williams et al. (2012) found long-lasting improvements in 'mentalising'. In terms of affective inferencing during reading, ToM has been investigated with certain language abilities (Abbeduto et al., 2004; Astington & Jenkins, 1999; Pluck et al., 2021), but not with emotional inferencing. Therefore, we will investigate the potential positive relationship between ToM and affective inferencing abilities.

Inhibitory control

IC denotes the voluntary regulation of goal-irrelevant stimuli, behaviours, and cognition (Diamond, 2013). As outlined by Kintsch (1998), when reading, we will inevitably form multiple representations of the text based on the variety of information from the text

itself and our own knowledge base. In building a coherent and context-specific mental model, readers must strengthen context-appropriate information and suppress irrelevant ones, which is where IC contributes. In terms of inferencing, this also means inhibiting inappropriate inferences (Cain, 2006; Kieffer et al., 2013; Kintsch, 1998).

Studies have shown IC's involvement in inferencing and reading comprehension. Nakamichi (2011) found that response inhibition was positively correlated to students' ability to make conditional inferences (i.e. drawing conclusions from if-then propositions). Additionally, performance in inferential questions was positively linked to inhibition for 8- to 15-year-olds (Potocki et al., 2017). Kieffer et al. (2013) reported that the relationship between IC and reading comprehension was uniquely direct. Indeed, less-skilled comprehenders were less able to inhibit irrelevant nouns in a sentence-completion task (Cain, 2006; De Beni & Palladino, 2000). Gernsbacher and Faust (1991) reached the same conclusion in tasks requiring suppression of inappropriate meanings of ambiguous words or homophones. Hence, the link between IC with conditional inference and other forms of discourse comprehension suggests that IC could predict emotional inferencing abilities.

Linguistic factors

To fully investigate the contributions of ToM and IC in emotional inferencing, three linguistic factors will be controlled. Print exposure has been linked to language abilities (e.g. Kim et al., 2021; Lee et al., 2019) from infancy to early adulthood (Mol & Bus, 2011). In adolescents, reading fiction books contributed to inference-making in fiction passages (Duncan et al., 2016). Further, syllogistic reasoning, logical inferences made from two propositions (Bucciarelli & Johnson-Laird, 1999), has been significantly related to university students' print exposure (Siddiqui et al., 1998). Other linguistic factors such as vocabulary knowledge and reading enjoyment have also been linked to inference abilities (Clinton, 2015;

Daugaard et al., 2017). As elaborated by Mol and Bus (2011) and West et al. (1993), reading more (i.e. exposure to print) increases language skills (e.g. vocabulary knowledge), which increases one's reading enjoyment. This in turn increases reading, which increases language skills, et cetera. Therefore, due to their cyclical relationship with each other and their relationships with inference abilities, print exposure, vocabulary knowledge and reading enjoyment will be controlled to explore the independent effects of ToM and IC.

Studying adolescents

While few studies have focused on inference-abilities in adolescents, higher-level skills such as IC seem to become more important as readers age and reading-content difficulty increases (Chang, 2020). Blair et al. (2007) found stronger correlations between IC with phonemic awareness and letter knowledge for 6-year-old than 5-year-old students. Potocki et al. (2017) found that linguistic factors such as vocabulary and syntactic knowledge contributed to reading comprehension difficulties for students in Grades 3 and 5, but not 7 or 9. Instead, IC played a more significant role for the older students, even when linguistic factors were controlled. Therefore, IC may contribute more to emotional inferencing abilities in adolescents rather than in younger students.

With ToM, although early acquisition of this ability is important (Korkmaz, 2011), 'mentalising' skills develop through middle childhood (Devine & Hughes, 2013) and into early adulthood (Dumontheil et al., 2010; Valle et al., 2015). This development could be due to the increasingly complicated emotional environment and processing in adolescence. During this time, interpersonal relationships become more important than in previous years (Kenny et al., 2013) and significant changes occur in socio-emotional domains (Eccles et al., 2003). Perhaps relatedly, 'mentalising' abilities become more precise and complex (Valle et al., 2015). Considering this development and the more complex nature of narrative texts older

students must engage with, ToM could prove to be an important contributor to their ability to make affective inferences.

The present study

In summary, our study aims to investigate the relationship between ToM and IC with adolescents' individual differences in emotional inferencing abilities, quantified by reading time. We predict that adolescents with better ToM and IC will form stronger mental models for a character's emotional state implied by context. Hence, they will read congruent sentences faster and incongruent sentences slower than adolescents with poorer abilities, reflecting the better-established representation of the correctly implied emotion and thus better emotional-inferencing abilities. This effect of ToM and IC will remain when linguistic factors are controlled.

Methods

Participants

Two-hundred and ninety-six volunteer sixth-form students took part in the study; 173 participants (63.4% female, 34.4% male, 1.8% non-binary/third gender) were included in analysis, unless otherwise stated. Inclusion criteria included 16-18 years old, English as a first language, and no reported reading difficulties such as dyslexia. 16-18 years is the average age range of sixth-form students in the UK, and 16-year-olds and above are independently able to provide consent (British Psychological Society, 2021). The students attended grammar schools across London and the West Midlands. This study was approved by the University of Birmingham Ethical Guidelines Committee.

Cognitive and Linguistic Factors: Measures

Table 1 highlights the questionnaires and scales used to measure the cognitive and linguistic factors in the order presented to the participants.

Table 1

Details of Each Measure of Cognitive and Linguistic Factors Used in the Study

Measure	Description	Examples	Appendix
Reading habits questionnaire	Participants rate how much they enjoy reading on a 5-point Likert scale, 1 = *Really do not enjoy* to 5 = *Really enjoy reading*. The higher the score (maximum 5 points), the more their reading enjoyment.	"How much do you enjoy reading?"	*N/A*
Author Recognition Test (ART)	Print exposure test comprising 100 names, 50 authors and 50 non-authors. Participants select the names they are certain are authors. Correct selection adds 1 point, incorrect deducts 1 point for a maximum score of 50. Higher scores indicate more print exposure.	Authors: *"Haruki Murakami"* *"Cassandra Clare"* Non-authors: *"Mahmoud Akhtar"* *"Nadine Breen"*	*A*
Synonym-Antonym Test (S/A)	A 30-item vocabulary knowledge test, 15 synonyms and 15 antonyms. Participants choose the correct answer from 4 options; each correct answer adds 1 point. The maximum score is 30, with higher scores signifying better vocabulary knowledge.	Synonym: "*Help*" "Aid" "Assist" "Payment" "Application"	*B, C*
Impulse Control scale of the Self-Rated Executive Function Index (S-EFI/IC; Spinella, 2005)	5-item self-report questionnaire measuring inhibitory control. Participants rate how much they agree to a statement on a 5-point Likert scale (1 = *Strongly disagree* to 5 = *Strongly agree*). Scores are inverted and summed, with a maximum score of 25. Higher scores denote better inhibitory control. The IC scale has good internal consistency (α=.69) and convergent validity (Spinella, 2005). S-EFI measures satisfactory variance in healthy, non-clinical populations (Kruger, 2011).	"I take risks, sometimes for fun" "I lose my temper when I get upset"	*D*

Perspective Taking factor of the Questionnaire of Cognitive and Affective Empathy (QCAE/PT; Reniers et al., 2011)	Theory of mind test containing 10 questions. Participants indicate how much they agree with a statement on a 4-point Likert scale (1 = *Strongly disagree* to 4 = *Strongly agree*). The maximum score is 40, and higher scores signify better theory of mind. The PT factor has high internal consistency (α=.85) (Gomez et al., 2020) and test-retest reliability (0.76) (Liang et al., 2019). QCAE has good construct, convergent, and discriminative validities (Reniers et al., 2011; Liang et al., 2019).	"I can easily tell if someone else wants to enter a conversation." "I am good at predicting what someone will feel."	*E*

Emotional stimuli

We designed 40 emotional stimuli (see *Appendix F*) which suggested an emotion felt by a character, then revealed that the character feels an emotion congruent or incongruent to that suggested emotion. Stimuli were 2-sentence stories of characters experiencing emotion-evoking events. Emotions included basic and broader emotions (Fontaine & Scherer, 2013), with half of the stimuli comprising positively-valenced affects. Ten yes/no comprehension questions measured attention and understanding.

Each stimulus contained 5 regions (see *Table 2*). The first provides context, containing the first sentence of the stimulus and the narrative from which readers will infer an emotion. For example, *'Jack's dog died the other day,'* prompts the inferencing of the emotion 'sadness'. Region 2 contains the first part of the second sentence.

Region 3 is the single-word 'target' region, the first region where the congruence effect will be observed. This comprises either an explicit mention of emotion (e.g. *'jolly'*) or behaviour suggesting an emotion (e.g. *'beamed'*). For each narrative, two versions of the target word were made: congruent with context and incongruent. Two examples out of the 20 congruent-incongruent pairs can be seen in *Table 2*. Emotion words within pairs were matched on frequency ($t(19)= 1.915$; $p = .071$), based on the Subtlex-UK Word Frequency Database (van Heuven et al., 2014), and length, based on the number of letters ($t(19)= .737$; $p = .470$). Target words could only be repeated once, placed in one congruent and one incongruent stimulus, to ensure that participants only encountered them once due to counterbalancing (see below).

The fourth region measures the 'spillover' effects of congruence, as the effect typically extends the target word (Poynor & Morris, 2003; Tokowicz & Warren, 2010). It

contains the subsequent word or two words following the target. Finally, Region 5 contains the remainder of the second sentence.

We counterbalanced stimuli presentation to control for order and sequence effects. Two lists were made, each containing 20 different stimuli, to ensure that each participant only encountered a story once (i.e. either the congruent or incongruent version), preventing familiarity effects.

Table 2

Examples of Positive- and Negatively-Valenced Congruent Stimuli and Their Incongruent Counterparts, as well as the Different Regions Within Them

Congruent (*negatively valenced*)
Region 1: Jack's dog died the other day. (*context*)
Region 2: He
Region 3: grieved (*target*)
Region 4: the whole (*spillover*)
Region 5: week after.
Incongruent
R1: Jack's dog died the other day. (*context*)
R2: He
R3: grinned (*target*)
R4: the whole (*spillover*)
R5: week after.
Congruent (*positively valenced*)
R1: Vincent watched his favourite pianist perform beautifully. (*context*)
R2: He
R3: beamed (*target*)
R4: as the (*spillover*)
R5: pianist bowed at the end of the performance.
Incongruent
R1: Vincent watched his favourite pianist perform beautifully. (*context*)
R2: He
R3: booed (*target*)
R4: as the (*spillover*)
R5: pianist bowed at the end of the performance.

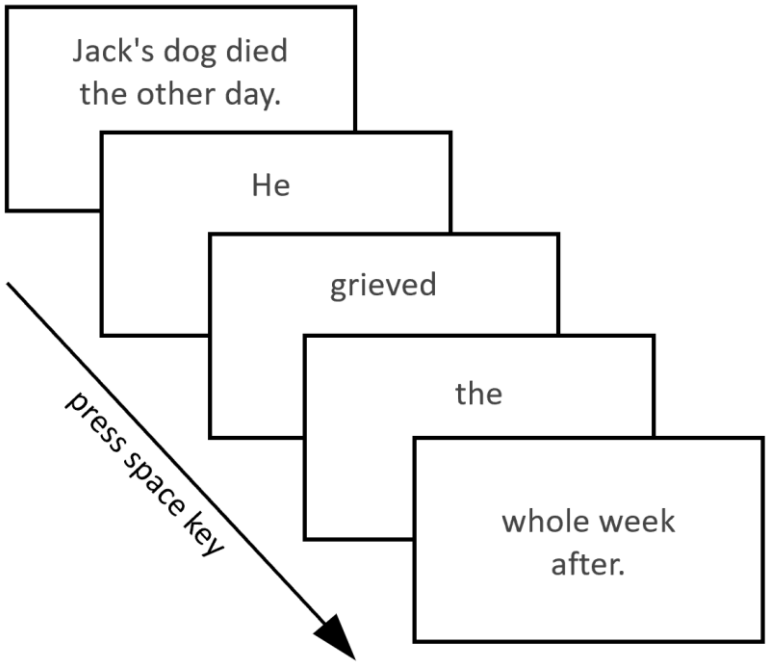
Procedure

Students completed the experiment as a group in their school computer labs under the supervision of their teachers. Students from one school completed it unsupervised. The experiment was split into two parts: a survey and a self-paced reading task. Participants received a Qualtrics (2020) survey containing the experiment information, consent form, demographics questions and measures of cognitive and linguistic factors in the order they were introduced above. Upon completion, students randomly received one of two links directing them to a self-paced reading task. Each link led to a different version of the task based on counterbalancing.

The self-paced reading task was run through Pavlovia using PsychoPy (Peirce et al., 2019). Instructions informed students to read the text at their normal reading pace and avoid distractions. Following this were 4 practice trials. The stimuli were presented one region at a time; pressing the spacebar moved onto the next region (see *Figure 3*). Furthermore, students answered comprehension questions using the ‘y/n’ keys for ‘yes/no’ respectively. Emotional-inference stimuli were interspersed among 60 distractor stimuli. The experiment terminated after 80 trials.

Figure 3.

Self-Paced Reading Procedure



Results

The procedure used for reading-time data removal during cleaning (41.6% removed) is as follows: (1) reading times below 200 (Jegerski, 2013) and above 2000 msec; (2) data points 2 standard deviations below or above the mean; (3) participants with less than 5 out of 10 correct comprehension-question answers. We calculated reading time differences for the target and spillover regions by subtracting the congruent from the incongruent reading times. Descriptive statistics are illustrated in *Table 3*.

Table 3

Means (M) and Standard Deviations (SD) of Reading Times and Scores of All Administered Measures

Measure			*M*	*SD*
Reading time (msec)	Target region	Congruent	521	113.3
		Incongruent	527	125
		Difference	6.4	68.0
	Spillover region	Congruent	522	128
		Incongruent	529	126
		Difference	7.2	64.5
Cognitive factors		QCAE/PT	31.61	4.47
		S-EFI/IC	16.13	3.33
Linguistic factors		ART	16.25	8.24
		S/A	21.75	3.55
		Reading enjoyment	3.41	1.05

Note. QCAE/PT = QCAE Perspective Taking score; S-EFI/IC = S-EFI Inhibitory Control score; ART = Author Recognition Test score; S/A = Synonym-Antonym test score

Spearman's correlations, investigating the relationship between reading time and all administered measures (see *Table 4*), were conducted due to the slight skew of certain measures. The three linguistic factors were significantly positively correlated; the higher a student's print exposure (ART), the more their vocabulary knowledge (S/A) and reading enjoyment. However, they were not correlated with reading-time differences. Interestingly,

negative correlations were found in the relationships between ToM (QCAE/PT) and vocabulary knowledge, as well as inhibitory control (S-EFI/IC) and reading enjoyment.

ToM was negatively correlated with reading time differences in the target region. Additional Spearman's analysis revealed a significant negative correlation between QCAE/PT score and incongruent-stimuli reading times in the target region ($r(171)= -.161$, $p= .034$), indicating that those with better 'mentalising' abilities read incongruent sentences faster than those with poorer ToM. However, ToM was not significantly correlated to congruent target-region reading-times. Furthermore, IC was not correlated with reading-time differences.

Table 4

Correlations Between the Reading-Time Differences, as well as Cognitive and Linguistic Factor Scores

Measure	1	2	3	4	5	6	7
1. Diff target	-						
2. Diff spillover	.426**	-					
3. QCAE/PT	-.198**	-.072	-				
4. S-EFI/IC	-.040	-.059	.079	-			
5. ART	-.017	-.075	.007	-.000	-		
6. S/A	-.032	.058	-.166**	-.126	.437**	-	
7. Enjoyment	-.072	.037	.075	-.183**	.403**	.271**	-

*Significant at $p < .05$; ** Significant at $p < .01$ (1-tailed)
Note. *df* = 220 for measures 1 and 2, *df* =171 for measures 3-7. Diff target = target region reading time difference; Diff spillover = spillover region reading time difference; QCAE/PT = QCAE Perspective Taking score; S-EFI/IC = S-EFI Inhibitory Control score; ART = Author Recognition Test score; S/A = Synonym-Antonym test score; Enjoyment = reading enjoyment score

Two hierarchical multiple regressions (target and spillover) were run to determine the amount of reading-time variance that can be explained by ToM and IC. Independent and controlled variables were entered in three steps: ART, S/A and Reading Enjoyment scores were entered in Step 1; QCAE/PT scores were added in Step 2; and S-EFI/IC were added in

the last step to measure the effects of both independent variables separately. Reading time difference was set as the dependent variable.

As seen in *Table 5*, only ToM was a significant predictor of target-region reading time differences, accounting for 4.4% of the variance. Together, all variables accounted for 5.9% of the variance in target-region reading time differences. No variable significantly predicted spillover-region reading-time differences.

Table 5

Results of Hierarchical Multiple Regression Analyses for the Target and Spillover Regions.

Model	Variable entered	Cumulative R^2	R^2 change	Beta
a. Dependent variable: target region reading-time difference				
1	ART	.008	.008	0.026 (p=.776)
	S/A			0.008 (p=.924)
	Read enjoyment			-0.095 (p=.258)
2	ART	.051	.044	0.027 (p=.759)
	S/A			-0.041 (p=.639)
	Read enjoyment			-0.072 (p=.384)
	QCAE/PT			-0.214 (p=.006)**
3	ART	.059	.008	0.042 (p=.634)
	S/A			-0.048 (p=.581)
	Read enjoyment			-0.101 (p=.243)
	QCAE/PT			-0.213 (p=.006)**
	S-EFI/IC			-0.093 (p=.240)
b. Dependent variable: spillover region reading-time difference				
1	ART	.012	.012	-.0124 (p=.170)
	S/A			0.069 (p=.429)
	Read enjoyment			0.054 (p=.516)
2	ART	.014	.001	-0.123 (p=.172)
	S/A			0.060 (p=.497)
	Read enjoyment			0.058 (p=.490)
	QCAE/PT			-0.037 (p=.642)
3	ART	.017	.003	-0.113 (p=.214)
	S/A			0.055 (p=.533)
	Read enjoyment			0.040 (p=.653)
	QCAE/PT			-0.036 (p=.649)
	S-EFI/IC			-0.060 (p=.456)

**Significant at $p < .01$

Note. QCAE/PT = QCAE Perspective Taking score; S-EFI/IC = S-EFI Inhibitory Control score; ART = Author Recognition Test score; S/A = Synonym-Antonym test score

Participant and item analyses were conducted to investigate the differences in congruent and incongruent reading times within the target and spillover regions. Furthermore, the potential effect of stimulus affective valence was investigated, comparing reading times between positive- and negative-valence items. Overall, 6 paired-samples t-tests were carried out (see *Table 6)*; means and standard deviations are shown in *Table 7*. Regarding participant and item analyses, while the results went in the expected direction, they did not reach full significance for either the target or the spillover regions. Additionally, there was no significant effect of stimulus affective valence on reading-time differences.

Table 6

Paired-Samples T-Test Results Comparing Congruent and Incongruent Reading Times from the Target and Spillover Regions

Region	*df*	*t*	*One-sided p*	*Two-sided p*
a. Participant analysis				
Target	221	-.896	.186	.371
Spillover	221	-.323	.374	.747
b. Item analysis				
Target	19	-1.151	.132	.264
Spillover	19	-.956	.176	.351
c. Valence-effect analysis				
Target	221	.422	.337	.674
Spillover	221	-.049	.480	.961

Table 7

Reading-Times Means and Standard Deviations from Each Analysis

Region	Condition	Reading time (msec)	
		M	SD
a. Participant analysis			
Target	Congruent	536.57	136.21
	Incongruent	541.73	143.94
Spillover	Congruent	547.71	343.40
	Incongruent	555.58	211.15
b. Item analysis			
Target	Congruent	519.34	59.87
	Incongruent	525.31	62.27
Spillover	Congruent	521.54	67.64
	Incongruent	528.55	63.29
c. Valence-effect analysis			
Target	Difference positive	7.62	93.34
	Difference negative	2.84	121.09
Spillover	Difference positive	6.81	94.65
	Difference negative	7.40	123.92

Note. Reading time differences for the valence-effect analyses were calculated by subtracting the incongruent from the congruent reading times.

Discussion

We hypothesised that ToM and IC will each uniquely predict individual differences in adolescent emotional inferencing abilities. In other words, those with better abilities will read affective sentences congruent to context faster and incongruent sentences slower than adolescents with poorer abilities. Furthermore, these effects will be independent of linguistic factors. Results partially supported these predictions.

Theory of mind

Correlational and regression analyses showed that ToM was a significant predictor of reading time differences in stimuli target regions. However, this was not the case for the spillover region. Furthermore, the direction of this significant relationship was unexpected; adolescents with better ToM read incongruent sentences faster than those with poorer ToM. While this opposes the notion that ToM is involved in forming mental models, potential alternative reasoning is that, instead of directly aiding the formation of more established and rigid models, ToM facilitates updating situational models. Vega (1996) showed that when an emotionally-incongruent sentence is introduced, reading times slow down. However, readers quickly update their mental model, demonstrated by faster reading times in subsequent sentences which remain incongruent. While this has not been investigated with ToM, ‘mentalising’ is heavily linked to cognitive flexibility (Jacques & Zelazo, 2005). Take, for example, the ability to flexibly adopt the perspective of another person which might contrast your own. Perhaps, adolescents with better ToM can more quickly and flexibly update their mental models, while those with poorer ToM take more time such that their reading time is slower in the target region. Nevertheless, due to the relatively small variance ToM explained (4.4%) and our study limitations (elaborated below), we cannot conclusively say that ToM is a significant or relevant enough predictor to be involved in reading-intervention research

comprising narrative literature with shifting emotions or perspectives. Therefore, ToM's role in updating mental models when reading affective texts should be investigated further.

This can be done using word probes instead of a self-paced reading task. Radvansky and Copeland (2010) investigated readers' abilities to update their mental model of spatial and temporal information. After the information shift, participants must respond yes/no to whether a probe occurred earlier in the story. This probe is either consistent with the updated information or with information prior to changes. The authors found differences in response times and accuracy to the probe (showing that readers updated their mental model) but not reading time differences. This may further support the use of word probes in place of reading time measures. In another study, participants read a sentence where an object's shape is implied and then another sentence where its shape changes. Following that, they were probed with a shape in line with the changed shape or the original (Hoeben Mannaert et al., 2019). Therefore, in terms of our stimuli, word probes will be either consistent with the congruent or incongruent emotion. If adolescents with better ToM have lower response times and higher accuracy than those with poorer ToM for probes consistent with the updated emotion, this would support the idea that ToM is involved in updating affective mental models.

Inhibitory control

Contrary to the hypothesis, IC was not correlated with or predictive of reading time differences. This is not consistent with previous findings reporting positive relationships between IC with inference abilities (Nakamichi, 2011) and reading comprehension (Kieffer et al., 2013), especially for adolescents (Potocki et al., 2017). Nonetheless, in a meta-analysis, Follmer (2018) reported mixed findings from studies investigating IC and adolescent reading comprehension. They associated this with the varying ways IC is conceptualised and measured, highlighting the need to focus on specific types of inhibition instead of IC as a

general construct. In a sample of adolescent students, Arrington et al. (2014) found that cognitive inhibition (the intentional suppression of task-irrelevant information and thoughts (Nigg, 2000)) was positively correlated with reading comprehension, but response inhibition (the deliberate suppression of automatic or prepotent responses to stimuli (Logan & Cowan, 1984)) was not. This is compatible with the idea that IC is involved in facilitating the formation of better mental models by inhibiting unnecessary information (Cain, 2006; Kintsch, 1998). The IC measure we used, on the other hand, likely measures response inhibition. Hence, individual differences in adolescent emotional inferencing should be investigated in relation to specific facets of IC such as cognitive inhibition.

Inter-variable correlations and limitations

Print exposure, vocabulary knowledge and reading enjoyment were positively correlated, congruous with previous works (Mol & Bus, 2011; West et al., 1993). Furthermore, these linguistic factors were not correlated to reading time differences. This agrees with Potocki et al.'s (2017) findings, that lower-level linguistic factors contribute less to reading comprehension for older than younger students.

Another relationship found was that ToM negatively correlated with Synonym-Antonym test scores; adolescents who were better at 'mentalising' had less vocabulary knowledge than those with better abilities. This was unexpected, as studies have found positive associations between ToM and verbal abilities in adolescents (Scheeren et al., 2013; Vetter et al., 2013). There may be a few reasons for this. In a population of 18- to 34-year-olds, Mainz et al. (2017) found that participant scores on a multiple-choice synonym test were *negatively* correlated to performance in a lexical decision task (LDT), a task measuring lexical activation and access. Although this relationship was no longer significant after significance level was adjusted, the authors highlighted that the relationship between

vocabulary and language processing might differ depending on the type of vocabulary test used. In another study, ToM of younger and older adults was *positively* correlated with LDT performance (Cavallini et al., 2013). Since LDT has been linked to abilities such as IC (Marí-Beffa et al., 2005; Passerieux et al., 1995), it is possible that the negative relationship we found between ToM and vocabulary knowledge was due to the mediating effect of IC. This provides an interesting direction for future research, as well as a methodological note about using a battery of vocabulary measures instead of one type (Mainz et al., 2017).

However, if this mediating effect is true, we should have found correlations between IC with both ToM (Meinhardt-Injac et al., 2020; Vetter et al., 2013) and vocabulary knowledge (Chung et al., 2020). However, this was not the case. Furthermore, we found that IC negatively correlated with reading enjoyment, another unexpected relationship. Together, these findings question the validity of the IC measure used. Although the Self-Rated Executive Function Index has been shown to demonstrate good construct and convergent validity (Spinella, 2005), it is possible that the Impulse Control scale particularly suffered from social desirability bias as adolescents are more susceptible to influences of social desirability than adults (Mwamwenda, 1995). Although it can be argued that this would affect all self-report questionnaires we used, the particularly strong, socially-undesirable nature of the questions (e.g. '*I make inappropriate sexual advances or flirtatious comments*') could exacerbate this effect (Krumpal, 2013). Therefore, we advise employing more discrete IC measures. Wonderlich et al. (2004), for example, selected and adapted items from self-report measures of trait impulsivity to best fit their sample and experimental outcome. Implicit IC measures could also be used (Arrington et al., 2014) to further avoid social-desirability effects.

Additional limitations

Although our measure of reading enjoyment resulted in a normal distribution of scores throughout the sample and significant correlations with other linguistic variables, using a single item might not have accurately captured students' perceptions of reading. Progress in International Reading Literacy Study (2006), for example, utilised 5 items measuring reading attitudes (e.g. '*I enjoy reading*.') and interactions (e.g. '*I read only if I have to.*'). Hence, future studies should use a combination of items that measure different ways of quantifying reading enjoyment.

Another limitation concerns our stimuli. While stimulus affective valence did not have a significant effect on reading times, consistent with Gygax et al.'s (2003) findings, participant and item analyses revealed that congruence had no significant effect on reading times in both target and spillover regions. This could be due to three things. First, some students completed the experiment without supervision which could have increased careless responding (Bowling et al., 2016). However, we accounted for this by selecting only participants who scored at least 5/10 on the comprehension questions.

Second, all participants came from grammar schools; these are schools that select applicants based on their academic achievement. Consequently, their reading-comprehension abilities may not be representative of the general student population. Their academic attainment at age 16, for example, is higher than average (Gorard & Siddiqui, 2018). The lack of reading-time differences can potentially be explained by their above-average abilities to update their mental models. Thus, if our sample was above average in their updating abilities, including students from both selective and non-selective schools would be recommended for a more representative sample.

Third, we must entertain the possibility that our stimuli did not evoke strong emotional inferences. One difference between our stimuli with other emotional-inference research is that our narrative was short, with only one sentence providing affective context. Gernsbacher et al. (1992), for example, used 6-sentence context narratives. Nonetheless, Mumper and Gerrig (2021) showed that simple 1-sentence narratives were sufficient for readers to make emotional inferences. Therefore, whether our stimuli successfully allowed readers to infer an affective state must be confirmed, potentially using LDTs to measure the activation of the context-inferred emotion.

Conclusion

Our study suggests the possibility that ToM is implicated in adolescents' abilities to update their mental model while making emotional inferences, specifically when the affective state of a character changes. We also highlight the need to investigate specific facets of IC such as cognitive inhibition when examining aspects of reading comprehension involving the formation of situational models. However, further investigation is needed to verify these relationships and to examine whether they are relevant predictors of emotional-inferencing abilities. Future studies should consider methodological considerations involving the selection of measures and sample populations. Nonetheless, if the relationship and relevance are supported, ToM and cognitive inhibition can be included in reading interventions as cognitive factors contributing to adolescent narrative literature comprehension.

References

Abbeduto, L., Short-Meyerson, K., Benson, G., & Dolish, J. (2004). Relationship between theory of mind and language ability in children and adolescents with intellectual disability. *Journal of Intellectual Disability Research, 48*(2), 150-159. https://doi.org/10.1111/j.1365-2788.2004.00524.x

Arrington, C. N., Kulesz, P. A., Francis, D. J., Fletcher, J. M., & Barnes, M. A. (2014). The contribution of attentional control and working memory to reading comprehension and decoding. *Scientific Studies of Reading*, *18*(5), 325-346. https://doi.org/10.1080/10888438.2014.902461

Astington, J. W., & Jenkins, J. M. (1999). A longitudinal study of the relation between language and theory-of-mind development. *Developmental Psychology, 35*(5), 1311–1320. https://doi.org/10.1037/0012-1649.35.5.1311

Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development*, *78*(2), 647-663. https://doi.org/10.1111/j.1467-8624.2007.01019.x

Bowling, N. A., Huang, J. L., Bragg, C. B., Khazon, S., Liu, M., & Blackmore, C. E. (2016). Who cares and who is careless? Insufficient effort responding as a reflection of respondent personality. *Journal of Personality and Social Psychology, 111*(2), 218–229. https://doi.org/10.1037/pspp0000085

British Psychological Society. (2021). *BPS Code of Human Research Ethics*. https://www.bps.org.uk/sites/bps.org.uk/files/Policy/Policy%20-%20Files/BPS%20Code%20of%20Human%20Research%20Ethics.pdf

Bucciarelli, M., & Johnson-Laird, P. N. (1999). Strategies in syllogistic reasoning. *Cognitive Science, 23*(3), 247-303. https://doi.org/10.1207/s15516709cog2303_1

Cain, K. (2006). Individual differences in children's memory and reading comprehension: An investigation of semantic and inhibitory deficits. *Memory, 14*(5), 553-569. https://doi.org/10.1080/09658210600624481

Cavallini, E., Lecce, S., Bottiroli, S., Palladino, P., & Pagnin, A. (2013). Beyond False Belief: Theory of Mind in Young, Young-Old, and Old-Old Adults. *The International Journal of Aging and Human Development*, 76(3), 181–198. https://doi.org/10.2190/AG.76.3.a

Chang, I. (2020). Influences of executive function, language comprehension, and fluency on young children's reading comprehension. *Journal of Early Childhood Research, 18*(1), 44-57. https://doi.org/10.1177/1476718X19875768

Chung, K. K. H., Lam, C. B., & Leung, C. O. Y. (2020). Contributions of executive functioning to Chinese and English reading comprehension in Chinese adolescent readers with dyslexia. *Reading and Writing, 33*(7), 1721-1743. https://doi.org/10.1007/s11145-020-10049-x

Clinton, V. (2015). Examining associations between reading motivation and inference generation beyond reading comprehension skill. *Reading Psychology, 36*(6), 473-498. https://doi.org/10.1080/02702711.2014.892040

Cutting, A. L., & Dunn, J. (1999). Theory of mind, emotion understanding, language, and family background: Individual differences and interrelations. *Child Development, 70*(4), 853-865. https://doi.org/10.1111/1467-8624.00061

Daniel, S. S., Walsh, A. K., Goldston, D. B., Arnold, E. M., Reboussin, B. A., & Wood, F. B. (2006). Suicidality, school dropout, and reading problems among adolescents. *Journal of Learning Disabilities, 39*(6), 507-514. https://doi.org/10.1177/00222194060390060301

Daugaard, H. T., Cain, K., & Elbro, C. (2017). From words to text: Inference making mediates the role of vocabulary in children's reading comprehension. *Reading & Writing, 30*(8), 1773-1788. https://doi.org/10.1007/s11145-017-9752-2

De Beni, R., & Palladino, P. (2000). Intrusion errors in working memory tasks: Are they related to reading comprehension ability?. *Learning and Individual Differences, 12*(2), 131-143. https://doi.org/10.1016/S1041-6080(01)00033-4

Devine, R. T., & Hughes, C. (2013). Silent films and strange stories: Theory of mind, gender, and social experiences in middle childhood. *Child Development, 84*(3), 989-1003. https://doi.org/10.1111/cdev.12017

Diamond, A. (2013). Executive functions. *Annual Review of Psychology, 64,* 135-168. https://doi.org/10.1146/annurev-psych-113011-143750

Dumontheil, I., Apperly, I. A., & Blakemore, S.-J. (2010). Online usage of theory of mind continues to develop in late adolescence. *Developmental Science, 13*(2), 331-338. https://doi.org/10.1111/j.1467-7687.2009.00888.x

Duncan, L. G., McGeown, S. P., Griffiths, Y. M., Stothard, S. E., & Dobai, A. (2016). Adolescent reading skill and engagement with digital and traditional literacies as predictors of reading comprehension. *British Journal of Psychology, 107*(2), 209-238. https://doi.org/10.1111/bjop.12134

Eccles, J. S., Barber, B. L., Stone, M., & Hunt, J. (2003). Extracurricular activities and adolescent development. *Journal of Social Issues, 59*(4), 865-889. https://doi.org/10.1046/j.0022-4537.2003.00095.x

Follmer, D. J. (2018). Executive function and reading comprehension: A meta-analytic review. *Educational Psychologist, 53*(1), 42-60. https://doi.org/10.1080/00461520.2017.1309295

Fontaine, J. J. R., & Scherer, K. R. (2013). The global meaning structure of the emotion domain: Investigating the complementarity of multiple perspectives on meaning. In J. J. R. Fontaine, K. R. Scherer, & C. Soriano (Eds.), *Components of emotional meaning: A sourcebook* (pp. 106–125). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199592746.003.0008

Gernsbacher, M. A. & Faust, M. E. (1991). The Mechanism of Suppression. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17* (2), 245-262. https://doi.org/10.1037/0278-7393.17.2.245

Gernsbacher, M. A. (1997). Two decades of structure building. *Discourse Process, 23*(3), 265-304. https://doi.org/10.1080/01638539709544994

Gernsbacher, M. A., Goldsmith, H. H., & Robertson, R. R. W. (1992). Do readers mentally represent characters' emotional states? *Cognition and Emotion, 6*(2), 89-111. https://doi.org/10.1080/02699939208411061

Gillioz, C., Gygax, P., & Tapiero, I. (2012). Individual differences and emotional inferences during reading comprehension. *Canadian Journal of Experimental Psychology, 66*(4), 239-250. https://doi.org/10.1037/a0028625

Gomez, R., Brown, T., Watson, S., & Stavropoulos, V. (2022). Confirmatory factor analysis and exploratory structural equation modeling of the factor structure of the Questionnaire of Cognitive and Affective Empathy (QCAE). *PLoS One, 17*(2), e0261914. https://doi.org/10.1371/journal.pone.0261914

Gorard, S., & Siddiqui, N. (2018). Grammar schools in England: a new analysis of social segregation and academic outcomes. *British Journal of Sociology of Education, 39*(7), 909-924. https://doi.org/10.1080/01425692.2018.1443432

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101*(3), 371-395. https://doi.org/10.1037//0033-295x.101.3.371

Grazzani, I., Ornaghi, V., Conte, E., Pepe, A., & Caprin, C. (2018). The relation between emotion understanding and theory of mind in children aged 3 to 8: The key role of language. *Frontiers in Psychology, 9,* 724. https://doi.org/10.3389/fpsyg.2018.00724

Gygax, P., Oakhill, J., & Garnham, A. (2003). The representation of characters' emotional responses: Do readers infer specific emotions? *Cognition and Emotion, 17*(3), 413-428. https://doi.org/10.1080/02699930244000048

Hall, C., Vaughn, S., Barnes, M. A., Stewart, A. A., Austin, C. R., & Roberts, G. (2020). The effects of inference instruction on the reading comprehension of English learners with reading comprehension difficulties. *Remedial and Special Education, 41*(5), 259-270. https://doi.org/10.1177/0741932518824983

Hoeben Mannaert, L. N., Dijkstra, K., & Zwaan, R. A. (2019). How are mental simulations updated across sentences?. *Memory & Cognition, 47*(6), 1201-1214. https://doi.org/10.3758/s13421-019-00928-2

Jacques, S., & Zelazo, P. D. (2005). Language and the Development of Cognitive Flexibility: Implications for Theory of Mind. In J. W. Astington & J. A. Baird (Eds.), *Why language matters for theory of mind* (pp. 144–162). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195159912.003.0008

Jegerski, J. (2013). Self-paced reading. In Jegerski, J., & VanPatten, B. (Eds.), *Research methods in second language psycholinguistics* (1st ed., pp. 36-65). Routledge. https://doi.org/10.4324/9780203123430

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference and consciousness*. Cambridge University Press.

Kendeou, P., Bohn-Gettler, C., White, M. J., & Van Den Broek, P. (2008). Children's inference generation across different media. *Journal of Research in Reading, 31*(3), 259-272. https://doi.org/10.1111/j.1467-9817.2008.00370.x

Kenny, R., Dooley, B., & Fitzgerald, A. (2013). Interpersonal relationships and emotional distress in adolescence. *Journal of Adolescence*, *36*(2), 351-360. https://doi.org/10.1016/j.adolescence.2012.12.005

Kieffer, M. J., Vukovic, R. K., & Berry, D. (2013). Roles of attention shifting and inhibitory control in fourth-grade reading comprehension. *Reading Research Quarterly*, *48*(4), 333-348. https://doi.org/10.1080/02702711.2019.1635239

Kim, D., Lowder, M. W., & Choi, W. (2021). Effects of print exposure on an online lexical decision task: a direct replication using a web-based experimental procedure. *Frontiers in Psychology*, *12*, 710663. https://doi.org/10.3389/fpsyg.2021.710663

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*(2), 163-182. https://doi.org/10.1037/0033-295X.95.2.163

Kneepkens, E. W., & Zwaan, R. A. (1995). Emotions and literary text comprehension. *Poetics, 23*(1-2), 125-138. https://doi.org/10.1016/0304-422X(94)00021-W

Korkmaz, B. (2011). Theory of mind and neurodevelopmental disorders of childhood. *Pediatric Research, 69*(8), 101-108. https://doi.org/10.1203/PDR.0b013e318212c177

Kruger G. H. (2011). Executive functioning and positive psychological characteristics: a replication and extension. *Psychological Reports, 108*(2), 477–486. https://doi.org/10.2466/04.09.21.PR0.108.2.477-486

Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: a literature review. *Quality & Quantity, 47*(4), 2025-2047. https://doi.org/10.1007/s11135-011-9640-9

Lee, H., Seong, E., Choi, W., & Lowder, M. W. (2019). Development and assessment of the Korean author recognition test. *Quarterly Journal of Experimental Psychology, 72*(7), 1837-1846. https://doi.org/10.1177/1747021818814461

Liang, Y. S., Yang, H. X., Ma, Y. T., Lui, S., Cheung, E., Wang, Y., & Chan, R. (2019). Validation and extension of the Questionnaire of Cognitive and Affective Empathy in the Chinese setting. *PsyCh Journal, 8*(4), 439–448. https://doi.org/10.1002/pchj.281

Logan, G. D., & Cowan, W. B. (1984). On the ability to inhibit thought and action: A theory of an act of control. *Psychological Review, 91*(3), 295–327. https://doi.org/10.1037/0033-295X.91.3.295

Luke, N., & Banerjee, R. (2013). Differentiated associations between childhood maltreatment experiences and social understanding: A meta-analysis and systematic review. *Developmental Review, 33*(1), 1-28. https://doi.org/10.1016/j.dr.2012.10.001

Mainz, N., Shao, Z., Brysbaert, M., & Meyer, A. S. (2017). Vocabulary knowledge predicts lexical processing: Evidence from a group of participants with diverse educational backgrounds. *Frontiers in Psychology*, *8*, 1164. https://doi.org/10.3389/fpsyg.2017.01164

Marí-Beffa, P., Hayes, A. E., Machado, L., & Hindle, J. V. (2005). Lack of inhibition in Parkinson's disease: evidence from a lexical decision task. *Neuropsychologia*, *43*(4), 638-646. https://doi.org/10.1016/j.neuropsychologia.2004.07.006

McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review, 99*(3), 440–466. https://doi.org/10.1037/0033-295X.99.3.440

Meinhardt-Injac, B., Daum, M. M., & Meinhardt, G. (2020). Theory of mind development from adolescence to adulthood: Testing the two-component model. *British Journal of Developmental Psychology*, *38*(2), 289-303. https://doi.org/10.1002/icd.298

Mol, S. E., & Bus, A. G. (2011). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin, 137*(2), 267–296. https://doi.org/10.1037/a0021890

Mumper, M. L., & Gerrig, R. J. (2021). The Representation of Emotion Inferences. *Discourse Processes*, *58*(8), 681-702. https://doi.org/10.1080/0163853X.2021.1882196

Mwamwenda, T. S. (1995). Age differences in social desirability. *Psychological Reports*, *76*(3), 825-826. https://doi.org/10.2466/pr0.1995.76.3.825

Nakamichi, K. (2011). Age differences in the relationship among conditional inference, working memory and prepotent response inhibition. *Psychologia*, *54*(1), 52-66. https://doi.org/10.2117/psysoc.2011.52

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin, 126*(2), 220–246. https://doi.org/10.1037/0033-2909.126.2.220

Oakhill, J. V., & Cain, K. (2012). The precursors of reading ability in young readers: evidence from a four-year longitudinal study. *Scientific Studies of Reading*, *16*(2), 91-121. https://doi.org/10.1080/10888438.2010.529219

Ong, D. C., Zaki, J., & Goodman, N. D. (2018). Computational models of emotion inference in theory of mind: A review and roadmap. *Topics in Cognitive Science, 11*(2), 338-357. https://doi.org/10.1111/tops.12371

Passerieux, C., Hardy-Bayle, M. C., & Widlöcher, D. (1995). Semantic inhibition disorders in schizophrenic patients assessed by a lexical decision task. *European Psychiatry*, *10*(1), 36-43. https://doi.org/10.1016/0767-399X(96)80073-3

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., Lindeløv, J. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods.* https://doi.org/10.3758/s13428-018-01193-y

Pluck, G., Córdova, M. A., Bock, C., Chalen, I., & Trueba, A. F. (2021). Socio-economic status, executive functions, and theory of mind ability in adolescents: Relationships with language ability and cortisol. *British Journal of Developmental Psychology*, *39*(1), 19-38. https://doi.org/10.1111/bjdp.12354

Potocki, A., Sanchez, M., Ecalle, J., & Magnan, A. (2017). Linguistic and cognitive profiles of 8- to 15-year-old children with specific reading comprehension difficulties: The role of executive functions. *Journal of Learning Disabilities*, *50*(2), 128-142. https://doi.org/10.1177/0022219415613080

Poynor, D. V., & Morris, R. K. (2003). Inferred goals in narratives: evidence from self-paced reading, recall, and eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(1), 3–9. https://doi.org/10.1037//0278-7393.29.1.3

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*(4), 515-526. https://doi.org/10.1017/S0140525X00076512

Progress in International Reading Literacy Study. (2006). *Progress in International Reading Literacy Study 2006 International Report*. https://timssandpirls.bc.edu/pirls2006/intl_rpt.html

Qualtrics. (2020, December). Qualtrics. Retrieved from https://www.qualtrics.com

Radvansky, G. A., & Copeland, D. E. (2010). Reading times and the detection of event shift processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(1), 210–216. https://doi.org/10.1037/a0017258

Reniers, R. L. E. P., Corcoran, R., Drake, R., Shryane, N. M., & Völlm, B. A. (2011). The QCAE: A questionnaire of cognitive and affective empathy. *Journal of Personality Assessment*, *93*(1), 84-95. https://doi.org/10.1080/00223891.2010.52848

Savolainen, H., Ahonen, T., Aro, M., Tolvanen, A., & Holopainen, L. (2008). Reading comprehension, word reading and spelling as predictors of school achievement and

choice of secondary education. *Learning and Instruction*, *18*(2), 201-210. https://doi.org/10.1016/j.learninstruc.2007.09.017

Scheeren, A. M., de Rosnay, M., Koot, H. M., & Begeer, S. (2013). Rethinking theory of mind in high-functioning autism spectrum disorder. *Journal of Child Psychology and Psychiatry*, *54*(6), 628-635. https://doi-org.ezproxye.bham.ac.uk/10.1111/jcpp.12007

Siddiqui, S., West, R. F., & Stanovich, K. E. (1998). The influence of print exposure on syllogistic reasoning and knowledge of mental-state verbs. *Scientific Studies of Reading*, *2*(1), 81-96. https://doi.org/10.1207/s1532799xssr0201_4

Snow, C. E., & Sweet, A. P. (2003). Reading for comprehension. In A. P. Sweet, & C. E. Snow (Eds.), *Rethinking reading comprehension*. The Guilford Press.

Spinella, M. (2005). Self-rated executive function: development of the executive function index. *International Journal of Neuroscience*, *115*(5), 649-667. https://doi.org/10.1080/0020745059052

Tokowicz, N., & Warren, T. (2010). Beginning adult L2 learners' sensitivity to morphosyntactic violations: A self-paced reading study. *European Journal of Cognitive Psychology*, *22*(7), 1092-1106. https://doi.org/10.1080/09541440903325178

Valle, A., Massaro, D., Castelli, I., & Marchetti, A. (2015). Theory of mind development in adolescence and early adulthood: The growing complexity of recursive thinking ability. *Europe's Journal of Psychology*, *11*(1), 112-124. https://doi.org/10.5964/ejop.v11i1.829

Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. Academic Press. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.473.5491&rep=rep1&type=pdf

van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). Subtlex-UK: A New and Improved Word Frequency Database for British English. *Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190. https://doi.org/10.1080/17470218.2013.850521

Vega, M. (1996). The representation of changing emotions in reading comprehension. *Cognition & Emotion, 10*(3), 303-322. https://doi.org/10.1080/026999396380268

Vetter, N. C., Altgassen, M., Phillips, L., Mahy, C. E., & Kliegel, M. (2013). Development of affective theory of mind across adolescence: disentangling the role of executive functions. *Developmental Neuropsychology*, *38*(2), 114-125. https://doi.org/10.1080/87565641.2012.733786

West, R. F., Stanovich, K. E., & Mitchell, H. R. (1993). Reading in the real world and its correlates. *Reading Research Quarterly*, 35-50. https://doi.org/10.2307/747815

Williams, B. T., Gray, K. M., & Tonge, B. J. (2012). Teaching emotion recognition skills to young children with autism: a randomised controlled trial of an emotion training programme. *Journal of Child Psychology and Psychiatry, 53*(12), 1268-1276. https://doi.org/10.1111/j.1469-7610.2012.02593.x

Wonderlich, S. A., Connolly, K. M., & Stice, E. (2004). Impulsivity as a risk factor for eating disorder behavior: Assessment implications with adolescents. *International Journal of Eating Disorders*, *36*(2), 172-182. https://doi.org/10.1002/eat.20033

Xavier, J., Tilmont, E., & Bonnot, O. (2013). Children's synchrony and rhythmicity in imitation of peers: Toward a developmental model of empathy. *Journal of Physiology Paris, 107*(4), 291-297. https://doi.org/10.1016/j.jphysparis.2013.03

Appendix A

Author Recognition Test Items

Authors	Non-authors
Virginia Woolf	Melissa Newcomb
J. B. Priestly	Julius Yairi
Ian Fleming	William Reynolds
Margaret Atwood	Mahmoud Akhtar
Kazuo Ishiguro	R.D. Lynas
Phillip Pullman	Geoffrey Llewellyn
Markus Zusak	Ben French
Robert Louis Stevenson	Nadine Breen
William Blake	Fern Elizabeth Rogers
Leo Tolstoy	Akemi Demura
Veronica Roth	Catherine Irving
Suzanne Collins	Andrew McQueen
Dan Brown	Luke Oaks
Ian McEwan	Carol Bateson
Arthur Miller	Blake Seymour
J.D. Salinger	Agnes Lacey
Jane Austen	Dmitry Abakumov
Louise Rennison	Tiffany Goodwin
J.K. Rowling	Joseph Bush
Ernest Hemingway	Elizabeth Barrows
Charlotte Brontë	Rudy Knight
Marcel Proust	Alice Gold
Harper Lee	Dean Skillbeck
William Golding	Bruno Jasper Allison
George Orwell	Anastasia Peterson
Anton Chekov	Hugo Lightfellow
James Patterson	Iain Elliott
George Elliott	Rashid Zafir
Carol Ann Duffy	Leopold Rutter
Terry Pratchett	Lazarus Kane

Sir Arthur Conan Doyle	Robert Addison
Lewis Carroll	Eilish Rownan
Wilfred Owen	Lazlo de Kuldes
Alice Sebold	Nore Yasagami
Stephen King	Lewis Ogden
Rudyard Kipling	Anton Gillespie
Malorie Blackman	Jane Andrew Forsyth
Michael Morpurgo	Sir Gerald Vineyard
Cormac McCarthy	Elizabeth Brown
Haruki Murakami	Akihiro Takeba
Roald Dahl	Clara Dains
Eoin Colfer	Michael Godwin
Thomas Keneally	George Ramona
Jonas Jonasson	Sam Pritchard
Paulo Coelho	Jeffrey Leavens
J.R.R. Tolkien	Andrew Blaze
Anthony Horowitz	William J. Ross
Cassandra Clare	Eliza Woodward
Mary Shelley	Fiore McCloud
George R Martin	Emily J. Johnson

Appendix B

Synonym Test

This appendix contains the instructions and items given to the participants for the Synonym Test.

Instruction: For the following 15 words, please click the answer that is a **synonym** of the word given. Only 1 answer per word is allowed. Please do not cheat as we will not be able to use your answers.
A synonym is a word that means (exactly or nearly) the same as another word.

Items:

1. Help
    a. Aid
    b. Assist
    c. Payment
    d. Application

2. Require
    a. Find
    b. Make
    c. Help
    d. Need

3. Advocate
    a. Pronounce
    b. Support
    c. Determine
    d. Deny

4. Ambivalent
    a. Equivocal
    b. Certain
    c. Affectionate
    d. Loathing

5. Cynical
    a. Optimistic
    b. Annular
    c. Sceptical
    d. Impolite

6. Acumen
    a. Rhythm
    b. Intelligence
    c. Honesty
    d. Memory

7. Camaraderie
    a. Secrecy
    b. Formality
    c. Judgement
    d. Fellowship

8. Abide
    a. Unite
    b. Tolerate
    c. Escape
    d. Maintain

9. Vocation
    a. Holiday
    b. Occupation
    c. Break-up
    d. Virtue

10. Essential
    a. Integral
    b. Useless
    c. Sharp
    d. Insignificant

11. Accomplish
    a. Expected
    b. Achieve
    c. Goal
    d. Purpose

12. Peril
    a. Safety
    b. Shelter
    c. Danger
    d. Mistrust

13. Omit
    a. Throw away
    b. Acquire
    c. Delete
    d. Add

14. Precious
    a. Necessary
    b. Valuable
    c. Irregular
    d. Simple

15. Calamitous
    a. Confident
    b. Dramatic
    c. Catastrophic
    d. Accidental

Appendix C

Antonym Test

This appendix contains the instructions and items given to the participants for the Synonym Test.

Instructions: For the 15 following words, please click the answer that is an **antonym** of the word given. Only 1 answer per word allowed. Please do not cheat as we will not be able to use your answers.
An antonym is a word that means the opposite to another word.

Items:

1. Concerned
    a. Passionate
    b. Ready
    c. Indifferent
    d. Stolen

2. Relegate
    a. Prevaricate
    b. Promote
    c. Import
    d. Keep

3. Anomaly
    a. Omen
    b. Possibility
    c. Regularity
    d. Tangent

4. Disparate
    a. Homogeneous
    b. Different
    c. Vast
    d. Despair

5. Abrupt
    a. Sudden
    b. Loud
    c. Gradual
    d. Regular

6. Vicious
    a. Malicious
    b. Brutal
    c. Gentle
    d. Fierce

7. Obscure
    a. Vague
    b. Ambiguous
    c. Clear
    d. Abstruse

8. Abundant
    a. Disadvantage
    b. Plentiful
    c. Ample
    d. Scarce

9. Divulge
    a. Disclose
    b. Promulgate
    c. Conceal
    d. Declare

10. Loathe
    a. Cherish
    b. Detest
    c. Prize
    d. Reject

11. Forgo
    a. Abstain
    b. Indulge
    c. Revere
    d. Refrain

12. Conquer
    a. Surrender
    b. Descend
    c. Punish
    d. Divide

13. Oppress
    a. Suppress
    b. Hurt
    c. Moan
    d. Permit

14. Adversity
    a. Misfortune
    b. Difficulty
    c. Adventurous
    d. Advantage

15. Methodical
    a. Disorganised
    b. Disobedient
    c. Disciplined
    d. Discipleship

Appendix D

Impulse Control scale of the Self-Rated Executive Function Index

This appendix contains the instructions and items given to the participants from the Impulse Control scale of the Self-reported Executive Function Index (S-EFI). Note that, in the survey given to the participants, the whole S-EFI was included with all its scales.

Instructions: This questionnaire contains 27 questions. Please rate each item using the scale below (strongly disagree to strongly agree).

Items:

I take risks, sometimes for fun.

Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
O	O	O	O	O

I do or say things that others find embarrassing.

Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
O	O	O	O	O

I make inappropriate sexual advances or flirtatious comments.

Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
O	O	O	O	O

I lose my temper when I get upset.

Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
O	O	O	O	O

I swear/use obscenities.

Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
O	O	O	O	O

Appendix E

Perspective Taking factor of the Questionnaire of Cognitive and Affective Empathy

This appendix contains the instructions and items given to the participants from the Perspective Taking factor of the Questionnaire of Cognitive and Affective Empathy. All item-responses replicate the options shown for Item 1 below.

Instructions: This questionnaire contains 10 questions. Please read each statement and select the option that best reflects the degree to which a particular statement relates (or does not relate) to you.

Items:

1. I can easily tell if someone else wants to enter a conversation.
    - O Strongly agree
    - O Slightly agree
    - O Slightly disagree
    - O Strongly disagree
2. I can pick up quickly if someone says one thing but means another.
3. I am good at predicting what someone will feel.
4. I am quick to spot when someone in a group is feeling awkward or uncomfortable.
5. Other people tell me I am good at understanding how they are feeling and what they are thinking.
6. I can easily tell if someone else is interested or bored with what I am saying.
7. I can sense if I am intruding, even if the other person does not tell me.
8. I can easily work out what another person might want to talk about.
9. I can tell if someone is masking their true emotion.
10. I am good at predicting what someone will do.

Appendix F

Self-Paced Reading Task Stimuli List

Valence	Congruent	Incongruent
Negative	Jack's dog died the other day. He grieved the whole week after.	Jack's dog died the other day. He grinned the whole week after.
	Lucas found out that someone had vandalised his painting. He exploded when he saw his hard work ruined.	Lucas found out that someone had vandalised his painting. He giggled when he saw his hard work ruined.
	Alison was watching a scary movie on Halloween. She screamed at the sudden noise.	Alison was watching a scary movie on Halloween. She cheered at the sudden noise.
	Mark had his job interview at the law firm last week. He trembled as he entered the interview room.	Mark had his job interview at the law firm last week. He chuckled as he entered the interview room.
	Monica watched as the school bully threw her bag out of the window. She yelled at the bully while her classmates stood behind her.	Monica watched as the school bully threw her bag out of the window. She beamed at the bully while her classmates stood behind her.
	Nick made a racist joke during a stand up comedy show. The crowd booed at him as he kept talking.	Nick made a racist joke during a stand up comedy show. The crowd clapped at him as he kept talking.
	Riley got a phone call to say their husband had a heart attack. They panicked as they heard the news.	Riley got a phone call to say their husband had a heart attack. They smiled as they heard the news.
	Gillian was trying to study for her exam in the library and the people sitting next to her started talking loudly. Gillian felt irritated as she flipped a page from her book.	Gillian was trying to study for her exam in the library and the people sitting next to her started talking loudly. Gillian felt contented as she flipped a page from her book.
	Olivia's brother kept knocking on her door during a video meeting. Olivia felt annoyed as she continued listening to the speaker.	Olivia's brother kept knocking on her door during a video meeting. Olivia felt hopeful as she continued listening to the speaker.
	None of Angellica's friends said happy birthday to her today. She felt disheartened as she walked back home.	None of Angellica's friends said happy birthday to her today. She felt heartened as she walked back home.
Positive	The lost man thanked Michael for showing the way. Michael grinned as he heard the gratitude.	The lost man thanked Michael for showing the way. Michael sobbed as he heard the gratitude.

	Vincent watched his favourite pianist perform beautifully. He beamed as the pianist bowed at the end of the performance.	Vincent watched his favourite pianist perform beautifully. He booed as the pianist bowed at the end of the performance.
	Jenny fainted on the road yesterday and some padders-by took her to the hospital. Jenny appreciated that they helped her.	Jenny fainted on the road yesterday and some padders-by took her to the hospital. Jenny complained that they helped her.
	Linda played well and her football team won the match. She applauded their success as her teammates all ran to hug her.	Linda played well and her football team won the match. She despised their success as her teammates all ran to hug her.
	Patricia’s kids helped her with Christmas decorations this morning. She kissed them after they all finished decorating.	Patricia’s kids helped her with Christmas decorations this morning. She punished them after they all finished decorating.
	Jenny’s husband bought her a bouquet of flowers. She cuddled him as he gave them to her.	Jenny’s husband bought her a bouquet of flowers. She scorned him as he gave them to her.
	Johnny got the highest grade in his class. He bragged all day at school.	Johnny got the highest grade in his class. He sulked all day at school.
	Steven Williams was a celebrated WWI soldier. They honoured him every Armistice Day.	Steven Williams was a celebrated WWI soldier. They blamed him every Armistice Day.
	Izzy was gardening when she heard her favourite song playing from her neighbour's house. She hummed as she pulled out some weeds.	Izzy was gardening when she heard her favourite song playing from her neighbour's house. She shuddered as she pulled out some weeds.
	Thomas got handed an 80% discount for food delivery. He felt jolly as he put the discount code into the delivery app.	Thomas got handed an 80% discount for food delivery. He felt anxious as he put the discount code into the delivery app.