Journal of Child Language (2026), 1–24
doi:10.1017/S0305000926100543

RESEARCH ARTICLE

Who gets it? Explaining variability in children's written irony comprehension

Henri Olkoniemi[1,2], Tuomo Häikiö[2], Matti Laine[3] and Penny M. Pexman[4]

[1]Unit of Psychology, University of Oulu, Finland; [2]Department of Psychology and Speech-Language Pathology, University of Turku, Finland; [3]Department of Psychology, Åbo Akademi University, Finland and [4]Department of Psychology, Western University, Canada

Corresponding author: Henri Olkoniemi; Email: henri.olkoniemi@oulu.fi

(Received 23 June 2025; revised 16 February 2026; accepted 20 February 2026)

Abstract

Understanding verbal irony involves detecting that the speaker's intended meaning contrasts with the literal meaning. This is challenging for children as the underlying skills required to understand irony may not be fully developed. We investigated how 10-year-olds' working memory, empathy skills, and gender were related to their processing and comprehension of written irony. Data from two previous eye-tracking experiments with 97 children (46 girls and 51 boys) were analysed. Results showed that children with stronger empathy skills had higher irony comprehension accuracy and were less likely to reread ironic phrases. Higher working memory was linked to faster processing of irony but did not lead to higher comprehension. Conversely, lower working memory was associated with more accurate irony comprehension. Child gender was not related to irony comprehension. These results imply that working memory and emotional perspective-taking are important for children's irony comprehension, underscoring theories that take individual differences into account.

Keywords: language development; irony; eye movements; working memory; empathy

Abstrakti

Ironisen kielen ymmärtäminen edellyttää, että kuulija havaitsee puhujan tarkoittaman merkityksen olevan ristiriidassa lauseen kirjaimellisen merkityksen kanssa. Tämä on lapsille haastavaa, sillä ironian tulkitsemiseen tarvittavat taidot eivät ole vielä täysin kehittyneet. Tässä tutkimuksessa selvitimme, miten 10-vuotiaiden lasten työmuisti, empatiataidot ja sukupuoli liittyivät kirjoitetun ironian prosessointiin ja ymmärtämiseen. Analysoimme kahden aiemman katseenseurantatutkimuksen aineiston, joka koostui 97 lapsesta (46 tyttöä ja 51 poikaa). Tulokset osoittivat, että lapsilla, joilla oli vahvemmat empatiataidot, oli parempi ironian ymmärtämisen tarkkuus, ja he palasivat harvemmin lukemaan ironisia ilmauksia uudelleen. Parempi työmuisti liittyi ironian nopeampaan prosessointiin, mutta ei parempaan ymmärtämiseen. Sen sijaan heikompi työmuisti oli yhteydessä tarkempaan ironian ymmärtämiseen. Lapsen sukupuoli ei ollut yhteydessä ironian ymmärtämiseen.

Tulokset viittaavat siihen, että työmuisti ja empatiakyky ovat tärkeitä tekijöitä lasten ironian ymmärtämisessä, ja ne ovat linjassa yksilöllisiä eroja huomioivien ironian ymmärtämisen teorioiden kanssa.

1. Introduction

In verbal irony, the literal meaning of the phrase and the intended meaning contradict each other (e.g. Attardo, 2000; Booth, 1974).[1] For example, in the animated Disney movie *Mulan* (1998), Chi-Fu, General Li, and Lee Shang are talking in a tent (Bancroft & Cook, 1998). The general promotes Lee Shang, and he brags about the new title: "Captain Lee Shang. Leader of China's finest troops. No, the greatest troops of all time." When they come out of the tent, chaos awaits them; the troops are fighting with each other while chickens are running around the yard. Chi-Fu says: "Most impressive." Chi-Fu clearly does not mean what he says. He is being ironic. Although Mulan is an animated movie aimed primarily at children, understanding the intended meaning – that Chi-Fu did not find the troops impressive – requires making complex inferences about the speaker's intent, which has been shown to be a challenge for children (e.g. Fuchs, 2023). The ability to comprehend irony develops late and requires the development of skills beyond those needed to understand literal language. In this paper, we combined data of two previous eye-tracking experiments investigating 10-year-old children's processing of written irony (Olkoniemi et al., 2023, 2025). By doing so, we created the opportunity to explore individual differences in the processing and comprehension of written irony, using measures that were collected but not analysed in the two previous studies due to their limited sample sizes. Before discussing these measures in detail, we first review what is known about the development of irony comprehension and the individual differences that have been proposed to influence its development.

1.1. Development of irony comprehension

Theoretical views. Most theories of irony comprehension do not explicitly account for developmental changes, despite efforts to integrate into theories the existing empirical findings on the development of irony comprehension (Creusere, 1999; Wilson, 2013). However, some theories of irony do consider individual differences and thus identify factors that children may need to develop to support their understanding of irony (e.g. Fabry, 2021; Pexman, 2008). For example, *the parallel-constraint-satisfaction framework* (Pexman, 2008) suggests that interpreters consider likely interpretations of the phrase (e.g. literal, a lie, or ironic) in parallel. The probability of each interpretation is mediated by contextual cues and interpreter-related factors, such as executive functions, and by the ability to understand others' emotions and mental states.

[1]There is debate about the definition of irony. We have given the classic definition here, but it has been argued that the classic definition does not capture the full range of irony use. For example, people use hyperbole to express ironic meaning, which this definition does not quite capture (Gibbs & Colston, 2007). However, there is currently no single definition upon which all scholars agree. The classical definition, though narrower, covers the types of irony included in the stories in the present paper. Thus, we offer the classic definition as it fits the focus of the paper.

https://doi.org/10.1017/S0305000926100543 Published online by Cambridge University Press

Another recent theory taking individual differences into account is the *predictive processing account* (Fabry, 2021). It is based on a broader assumption that living organisms make predictions about future states of the world to minimise surprise at their observations (*predictive coding theory*, Friston et al., 2006). The theory views communication as a two-way exchange of signals (e.g. speech or text) that produces predictions in both the communicator and the interpreter (Friston & Frith, 2015). In communication, irony use is often unexpected and is less frequent than literal language; thus, the priors are typically set to predict literal language (Fabry, 2021). When irony is encountered, a prediction error occurs, and a corrective process is required to comprehend the intended meaning. In addition, the priors may be adjusted to minimise prediction errors in future. This account also proposes that individual differences in perception, cognition, and emotional skills affect the ability to make predictions about the use of irony (Fabry, 2021). When the prediction fails, individual differences affect the ability to recover from the prediction error. As both theories suggest that cognitive and emotional abilities influence the ease of irony comprehension, the theories can be extended to predict that the development of these abilities in childhood should influence the development of irony comprehension.

Empirical evidence. For most children, irony comprehension starts to emerge around the age of 5–6 years (see Fuchs, 2023, for a review). By this age, children have developed sufficient language and social perspective-taking skills to begin to understand some of the intended meanings of ironic phrases. Ironic compliments (negatively valenced statements intended to convey a positive nonliteral message) are not used as often in communication, and children's comprehension of these remarks tends to develop later than that for ironic criticism (Fuchs, 2023), which can even be difficult for some adults (e.g. Tiv et al., 2020). In the early years, irony comprehension accuracy is low, and children tend to rely on literal interpretation (Ackerman, 1983; Dews et al., 1996; Harris & Pexman, 2003; Loukusa & Leinonen, 2008). Many studies have assessed children's irony comprehension using forced-choice metalinguistic questions (e.g. Harris & Pexman, 2003). However, the results of the studies involving other methodologies suggest that implicit sensitivity to ironic meaning may precede explicit comprehension (Climie & Pexman, 2008; Köder & Falkum, 2021; Loukusa & Leinonen, 2008). For example, Köder and Falkum (2021) demonstrated in their eye-tracking study that although 3-to-5-year-olds tended to misinterpret ironic utterances as literal, they showed some awareness of the intended negative attitude in irony in their gaze patterns. In that study, children listened to stories while viewing accompanying pictures. During the target phrase, they were shown happy and angry face emoticons, and even the 3-to-5-year-old children were found to look more at angry than happy emoticons when hearing ironic criticism. Moreover, in a study by Loukusa and Leinonen (2008), short ironic stories were read to 3–9-year-old children, and two open comprehension questions (i.e. what the protagonist meant and how the children knew that) were asked after each story. They demonstrated that even some 3-year-olds were able to understand the intended meaning of some of the stories, although average comprehension accuracy for children younger than 6–7-years was at the chance level.

It has been argued that the ability to understand irony develops in a sequence of skills: first, children start to comprehend the ironic speaker's belief (i.e. speaker does not believe what they literally say) and then the ironic speaker's intent (i.e. speaker wants their true belief to be recognised; Winner & Gardner, 1993). Early studies showed that younger children often misinterpreted irony as a lie (e.g. Demorest et al., 1984; Winner & Leekam, 1991). This was taken as evidence that children had developed the ability to understand

https://doi.org/10.1017/S0305000926100543 Published online by Cambridge University Press

the speaker's belief, but not yet the speaker's intent (Demorest et al., 1984; Winner & Leekam, 1991). Accumulating evidence, however, shows that even older children typically misinterpret irony literally, and not as a lie (e.g. Ackerman, 1983; Capelli et al., 1990; Dews et al., 1996; Garfinkel et al., 2024; Harris & Pexman, 2003; Köder & Falkum, 2021; Loukusa & Leinonen, 2008; Recchia et al., 2010). Some studies have found that children tend to develop simultaneously the understanding of the ironic speaker's belief and intent (e.g. Filippova & Astington, 2010; c.f., Pexman & Glenwright, 2007). Children have been shown to have above chance-level comprehension accuracy for judgements of both speaker belief and intent by the age of 7–9 for ironic criticism (Fuchs, 2023), and some studies report near-perfect accuracy for both judgements at this age (e.g. Pexman et al., 2006, Exp 2). Many of these studies have used separate forced-choice and/or open comprehension questions for the speaker's belief and intent (e.g. Filippova & Astington, 2010; Pexman et al., 2006). The findings of a further study by Köder and Falkum (2021) also suggest that children may have some level of understanding of ironic speakers' emotions prior to understanding speakers' belief. However, additional socio-communicative aspects, like appreciation of the humour and teasing functions of irony, seem to develop later in adolescence (Glenwright et al., 2017).

Although many studies have found that children reach near adult-level accuracy in irony comprehension by middle childhood (Fuchs, 2023), there is evidence that this is not universally the case (Barich et al., 2025; Capelli et al., 1990; Demorest et al., 1984; Olkoniemi et al., 2023, 2025; Zajączkowska & Abbot-Smith, 2020). For example, eye-tracking studies examining written irony comprehension have consistently reported near chance-level accuracy for 10–12-year-olds (Barich et al., 2025; Olkoniemi et al., 2023, 2025). In these studies, including those from which the present paper draws its data, irony comprehension was assessed using speaker-belief questions (e.g. in the Mulan example, "Did Chi-Fu think the troops were impressive?"). These findings contrast with previous research indicating that children of this age typically answer belief questions with high accuracy (Fuchs, 2023). The cause of these mixed findings is unclear, but one possibility is that the difference in accuracy scores is due to task demands.

Previous studies have tested children's irony comprehension using spoken stories (e.g. Ackerman, 1983; Capelli et al., 1990; Demorest et al., 1984; Loukusa & Leinonen, 2008), spoken stories with illustrations (e.g. Agostino et al., 2017; Filippova & Astington, 2010; Garfinkel et al., 2024; Winner & Leekam, 1991), puppet shows (Climie & Pexman, 2008; Harris & Pexman, 2003; Nicholson et al., 2013; Pexman & Glenwright, 2007), and short age-appropriate videos (Dews et al., 1996; Glenwright et al., 2017; Zajączkowska & Abbot-Smith, 2020) as materials. These methods likely include paralinguistic cues (e.g. tone of voice) that aid children's comprehension (e.g. Ackerman, 1983; Capelli et al., 1990; Pexman, 2008). Moreover, in many studies, the experimenter asked comprehension questions and also presented possible interpretive options to choose from, which were sometimes pictures or objects (e.g. Ackerman, 1983; Capelli et al., 1990; Climie & Pexman, 2008; Demorest et al., 1984; Dews et al., 1996; Glenwright et al., 2017; Loukusa & Leinonen, 2008; Pexman & Glenwright, 2007; Zajączkowska & Abbot-Smith, 2020). This is obviously necessary especially with younger children who cannot read, but it may have provided additional support to children in the comprehension process.

In contrast, in eye-tracking studies of children's comprehension of written irony, children read and answer questions without interaction with an experimenter or other people, and this may make comprehension more difficult. Further, reading may pose greater challenges for children than listening (e.g. Olkoniemi et al., 2023). These

https://doi.org/10.1017/S0305000926100543 Published online by Cambridge University Press

methodological differences may explain why eye-tracking studies of reading have tended to find that children were still developing their irony comprehension skills at 10 years of age.

Regardless of methodology, previous studies (e.g. Olkoniemi et al., 2023; Zajączkowska & Abbot-Smith, 2020) have demonstrated that there is considerable individual variability in the accuracy of children's irony comprehension in middle childhood and into adolescence (e.g. Glenwright et al., 2017). In other words, while irony remains challenging for many children, others exhibit near-ceiling levels of comprehension accuracy. Before describing the research on individual differences in irony comprehension, we first provide a more comprehensive overview of how children process written irony and engage with reading.

1.2. Processing of written irony

The advantage of using eye-tracking to study reading of verbal irony is that the method allows us to follow the processing of ironic meaning moment by moment at the millisecond level (Rayner, 2009). Previous studies have shown that it typically takes longer to read ironic phrases than literal phrases (e.g. Filik & Moxey, 2010; Olkoniemi & Kaakinen, 2021). When compared to reading of literal phrases, adults demonstrate increased subsequent rereading of ironic phrases (i.e. look-back to) and a higher likelihood of returning to the parts of the story context that make the phrase ironic (Olkoniemi & Kaakinen, 2021). This is thought to reflect time taken to attempt to integrate the ironic meaning with the context (which is considered an important step in irony comprehension; e.g. Grice, 1975). Indeed, later rereading has been shown to reflect conscious efforts to build a comprehensive representation of the text (Hyönä & Nurminen, 2010). This slower processing for irony is not static but rather is attenuated over the course of an experiment. This is referred to as the *trial effect* (Olkoniemi & Kaakinen, 2021) or the *early-late effect* (Spotorno & Noveck, 2014). For example, Spotorno and Noveck (2014) showed that adults spent more time reading ironic phrases and the immediately following sentences (i.e. the spillover regions, e.g. Rayner, 2009) than their literal counterparts, but this slower processing of irony vs literal phrases faded towards the end of the experiment. It seems that when irony is encountered repeatedly, readers start to predict the coming irony, which makes subsequent ironic phrases easier to process, with reading times more similar to those for literal phrases.

Studies of children's processing of written irony have focused on children who are aged 10 or older. This is due to the fact that the reading performance of 10-year-olds is expected to be like that of adults for the reading of literal language (see Blythe & Joseph, 2011, for a review of eye-movement research in developing readers). Though similar, children's reading is not yet fluent and is expected to be generally slower due to making more and longer fixations, shorter saccades, and more regressions (Blythe & Joseph, 2011).

Indeed, studies comparing children and adults in the processing of written irony have found surprisingly similar processing patterns (Barich et al., 2025; Olkoniemi et al., 2023). Like adults, children show increased rereading of the target phrase and of the critical context for irony (i.e. the context that makes the target phrase ironic; Barich et al., 2025; Olkoniemi et al., 2023, 2025), and children also show trial effects (Olkoniemi et al., 2023, 2025). However, Olkoniemi et al. (2023) found three main differences between children and adults. First, children's reading focused more on immediate reading (i.e. first-pass reading) and adults' more on later rereading (i.e. look-backs and look-froms). This is in

https://doi.org/10.1017/S0305000926100543 Published online by Cambridge University Press

line with the findings that when reading for comprehension, younger readers tend to spend more time on first-pass reading than adults (Kaakinen et al., 2015). Second, children with higher-than-average irony comprehension accuracy showed faster reading times, whereas the opposite was true for adults. Third, unlike adults, children's comprehension of ironic meaning in stories was not reflected in their reading times. That is, unlike adults, children's reading patterns were the same when the irony was correctly understood and when it was not. So far, individual differences in children's processing and comprehension of written irony have not been investigated. However, eye-tracking studies with adults have shown that several variables of individual differences are related to the time-course of processing written irony.

1.3. Individual differences related to processing and comprehending irony

Social skills. Development of social skills has been shown to be important for children to be able to understand the intended meaning of irony (Fuchs, 2023). In particular, socio-emotional skills are associated with irony comprehension: empathy (i.e. the ability to understand and share the feelings of others; Riess, 2017) and/or the ability to recognise emotions have been associated with higher irony comprehension accuracy in children (Agostino et al., 2017; Nicholson et al., 2013) and also with faster processing of irony in adults (Olkoniemi et al., 2019a, 2019b). Nicholson et al. (2013) found that 8- to 9-year-old children with strong empathy skills tended to make faster judgements of speakers' ironic intent and had higher irony comprehension accuracy. Additionally, Agostino et al. (2017) reported that children's and adolescents' (6- to 15-year-olds) irony comprehension was mediated by their ability to understand emotive communication (i.e. the ability to modulate emotional expression according to social display rules).

In eye-tracking studies on written irony, adults with better emotion recognition and emotion-naming skills had shorter reading times for ironic phrases (Olkoniemi et al., 2019a). Consistent with these findings, clinical studies have demonstrated that lesions in the ventromedial prefrontal cortex are associated with deficits in both empathy skills and irony comprehension (Zald & Andreotti, 2010).

Socio-cognitive skills, and especially theory of mind (the ability to consider the thoughts of others; e.g. Perner & Wimmer, 1985), have been linked to developing irony comprehension. However, empirical evidence for this link in typically developing children has been mixed (see Fuchs, 2023, for a review). While several studies have highlighted the importance of a more complex theory of mind reasoning (e.g. the ability to consider what one person thinks about another person's thoughts) for irony, other research has failed to find evidence for this relationship (Fuchs, 2023). Taken together, these findings suggest that for typically developing children, the ability to take the emotional perspective of others may be more important than socio-cognitive reasoning for irony comprehension.

Working memory. In addition to socio-emotional skills, successful irony comprehension requires integration of the ironic phrase with the preceding context (e.g. Grice, 1975). To be able to do this successfully, the interpreter needs to suppress the salient literal meaning and keep the relevant contextual information in an accessible state (e.g. Giora, 1999; Olkoniemi & Kaakinen, 2021). This puts a strain on the individual's working memory. It is therefore not surprising that children's irony comprehension is related to their working memory capacity (WMC; Godbee & Porter, 2013) and to their executive functioning in general (Caillies et al., 2014; Zajączkowska & Abbot-Smith, 2020). For

https://doi.org/10.1017/S0305000926100543 Published online by Cambridge University Press

example, Godbee and Porter (2013) studied sarcastic irony, metaphor, and simile comprehension in individuals with Williams syndrome between the ages of 5 and 43 years and compared their performance with that of age-matched controls. Williams syndrome is a neurodevelopmental disorder that causes, for example, mild to moderate intellectual impairment, yet relatively good general language skills and sociability (Godbee & Porter, 2013). For control participants, higher WMC was correlated with more accurate irony comprehension. However, this relationship was not found in the individuals with Williams syndrome, perhaps because their irony comprehension accuracy was at floor (Godbee & Porter, 2013).

Furthermore, studies with adults have shown that higher WMC is associated with earlier and faster processing of ironic meaning. First, people with higher WMC were found to make faster ironic interpretations after viewing dialogues containing an ironic comment (Antoniou & Milaki, 2021). Second, in eye-tracking studies of written irony, adults with higher WMC had longer first-pass reading times of ironic phrases (Kaakinen et al., 2014; Olkoniemi et al., 2016) and faster reading times of the spillover region (Kaakinen et al., 2014). In contrast, readers with lower WMC had increased later rereading times (i.e. look-backs) of the ironic phrase (Olkoniemi et al., 2016, 2019a). Additionally, high-WMC readers tend to be more selective in their reading of ironic texts than low-WMC readers, being more linear in their reading and focusing more on rereading interpretation-relevant parts of the text (i.e. only the context that makes the phrase ironic; Olkoniemi et al., 2024). There are a few eye-tracking studies that have not found relationships between adults' WMC and their processing of written irony (Olkoniemi et al., 2019b; Parola & Bosco, 2022). At least some of the studies that reported a null relationship had very short materials that may not have been sufficiently taxing on adults' working memory (Olkoniemi et al., 2019b; Parola & Bosco, 2022). As children's reading comprehension ability and working memory are still developing, it was expected that they would need the support of working memory to process written irony and thus that a relationship would be observed between WMC and measures of processing for written irony.

Gender. In addition to WMC and socio-emotional skills, child gender may be associated with the development of written irony comprehension. There is a robust female advantage in reading achievement from kindergarten age to early adulthood (Manu et al., 2023). In addition, the findings of previous studies have suggested that girls have better social skills than boys (e.g. Hajovsky et al., 2021; Proverbio, 2023). Given these gender differences, it is possible that girls would outperform boys in irony comprehension, especially when irony is in written form. However, the results of studies directly examining gender differences in children's irony comprehension are mixed (Garfinkel et al., 2024; Harris & Pexman, 2003; Recchia et al., 2010; Rothermich et al., 2020).

Rothermich et al. (2020) investigated the relationships of age and gender to irony comprehension in 8–12-year-old children and found that girls were more accurate than boys at classifying ironic comments as insincere. Similarly, in their study on adults watching videorecorded conversations containing different kinds of indirect forms of speech, Rothermich and Pell (2015) showed that female participants performed better in recognising the speaker's sarcastic intent than did male participants. In contrast, Harris and Pexman (2003) did not find gender differences in their study of 5- to 8-year-old children's ability to recognise and interpret the intent of ironic criticisms and compliments presented in puppet shows. In addition, Garfinkel et al. (2024) found no gender differences in the performance of 8-year-old children in an irony comprehension task, and this was also true for Recchia et al. (2010) who examined the use of irony in family

https://doi.org/10.1017/S0305000926100543 Published online by Cambridge University Press

conversations with 3–7-year-old children. Although studies reporting no gender differences in children's irony comprehension outnumber those showing a gender difference, the use of written irony in the present study could add difficulty to the task and may create conditions to observe gender differences.

1.4. The present study

The aim of the present study was to investigate how individual differences in empathy, WMC, and gender are related to 10-year-old children's processing and comprehension of written irony. To this end, we combined data from two previous eye-tracking studies. In addition to analysing individual differences, we controlled for children's technical reading (i.e. word reading accuracy and fluency) and reading comprehension skills as well as their relative age. We controlled for relative age because there is evidence that children born earlier in the year have advantages over those born later in the year, for example in academic achievement (e.g. McPhillips & Jordan-Black, 2009).

Based on the results of previous studies, we expected that children would find irony harder to comprehend and would be slower to process irony than literal language. Processing effects were expected to be observed as increased rereading of ironic target phrases and increased likelihood of returning to the context critical for ironic interpretation (e.g. Olkoniemi & Kaakinen, 2021). We also expected that children would show trial effects wherein their increased rereading of the ironic phrases (vs literal phrases) diminishes towards the end of the experimental session (Olkoniemi et al., 2023). We also expected that processing and comprehension of irony would be moderated by individual differences in empathy (e.g. Nicholson et al., 2013), WMC (e.g. Godbee & Porter, 2013), and gender (e.g. Rothermich et al., 2020). Specifically, we assumed that children with higher empathy, children with higher WMC, and children who were girls would tend to have more accurate and faster irony comprehension.

2. Methods

2.1. Participants

We combined data from two previous studies (Olkoniemi et al., 2023, 2025). In both studies, the processing of written irony was examined in 10-year-old children. Olkoniemi et al. (2023) compared the processing and comprehension of irony in children (n = 33) and adults (n = 30). Olkoniemi et al. (2025) conducted a training study testing only children (n = 72) in pre- and post-test phases. To optimise the compatibility of the combined datasets, only data from children in the Olkoniemi et al. (2023) study and only pre-test data from the Olkoniemi et al. (2025) study were used in the present analyses. This resulted in a total of 105 participants. We removed two participants with missing data on some of the individual differences measures and excluded six participants originally included in the Olkoniemi et al. (2025) study because they had only comprehended one ironic item in the pre-test. This resulted in a final sample of 97 10-year-old children (46 girls, 51 boys, M_{Age} = 124 months, SD_{Age} = 3.21 months). Children were in the fourth grade and came from three different schools in the Turku and Oulu areas. The children had received approximately three years and three months of formal reading instruction and were native Finnish speakers (the language used in both studies). The bilingual status of the children was not assessed in either study; however, the schools were situated in communities where over 90% of the population speaks Finnish. They had no known reading difficulties and had normal or corrected-to-normal vision.

https://doi.org/10.1017/S0305000926100543 Published online by Cambridge University Press

Children's parents signed a written informed consent form prior to the experiment, and verbal assent was asked from each child. Both studies were conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee for Human Sciences at the University of Turku, Finland.

2.2. Apparatus

Children's eye movements were recorded using EyeLink Portable Duo (in both experiments, $n = 57$) and EyeLink 1000 Plus (SR Research Ltd. Ontario, Canada) eye-trackers (40 participants in Olkoniemi et al., 2025 study). The eye-trackers used are comparable in their data quality, including precision and accuracy (Lyu et al., 2023; SR Research, 2025). With both eye-trackers, 500 Hz sampling frequency and 9-point calibration were used (successful calibration was determined by $M_{error} < 0.5°$ in visual angle, error at each point <1°). For the Portable Duo, the stimuli were presented on a 17.3" Asus ROG G752V laptop monitor, and participants were seated 60 cm from the screen. For the EyeLink 1000 Plus, the stimuli were presented on a 24" Asus VG248QE monitor, and participants were seated 90–92 cm from the screen. With both monitors, a refresh rate of 120 Hz and a resolution of 1920 x 1080 pixels were used. A chin-and-forehead rest stabilised the participant's head.

2.3. Experimental stories

In the Olkoniemi et al. (2023) study, the materials consisted of 26 experimental stories. For the Olkoniemi et al. (2025) study, this number was increased to 44 by adding 18 new, similar stories. In the latter study, the children first read half of the stories in the pre-test phase and then the other half in the post-test phase. Only the pre-test data were used in the present analyses. Each story had an ironic and a literal version, and each participant read only one of these versions. The presentation order of the story versions (literal versus ironic) was counterbalanced across participants, and the order of story presentation was randomised. An example story is shown in Table 1.

Table 1. An example of an experimental story and inference and text memory questions translated from Finnish

Region	Version	Content
Beginning		Josh and Theo are on their way home from school.
Critical context	Literal	Theo dropped his key while they were walking and Josh offers to help him look for it.
	Ironic	Theo dropped his key while they were walking, but Josh refuses to help him look for it.
Target phrase		"You're a good mate"
Spillover region		Theo says.
End		Soon, they will continue their journey home.
Inference question		Does Theo think Josh is a bad friend?
Text memory question		Has Theo lost his backpack?

https://doi.org/10.1017/S0305000926100543 Published online by Cambridge University Press

In both experiments, the stories were 4–5 sentences long and similarly structured (please see the original studies for more details). The stories began with 1–2 sentences providing background information and were the same for both versions. Next, a context sentence introduced information that made the following phrase either ironic or literal (i.e. critical context). This was the only sentence that differed between the versions: the target phrase itself remained the same. After the target phrase, a spillover region described who had uttered the target phrase. Each story concluded with a neutral sentence describing the story resolution. For both experiments, each story was followed by two questions: a text memory question and an inference question designed to gauge how well participants understood the meaning of the target phrase. The proportion of correct answers was calculated for each participant.

2.4. Measures

Technical reading skill. Technical reading skill was measured using the word fluency subtest of Lukilasse II (Häyrinen et al., 2013), in which the children read aloud as many words as possible from a list of 105 words within 120-second time limit. The test was scored for the total number of correctly read words (range 0–105 points). See Table 2 for descriptive statistics.

Reading comprehension. Reading comprehension was measured using the maze task (Ronimus et al., 2022). The task is comprised of 16 texts, each with 4 words missing. Based on the textual cues, participants chose the missing word from four options. The task correlates strongly with a standardised paper-based reading comprehension test (see Ronimus et al., 2022, for more details). The task was scored for the total number of correctly selected words (range 0–64 points). See Table 2 for descriptive statistics.

Empathy. Children's empathy skill was assessed using the Index of Empathy for Children and Adolescents (Bryant, 1982), translated into Finnish (Olkoniemi et al., 2023). This is a 22-item, paper-and-pencil self-report questionnaire, which was developed to assess affective empathy in children 6 years of age and older. The scale contains short claims (e.g. "It makes me sad to see a boy/girl who can't find anyone to play with"), for which a binary answer (yes/no) is given. Each empathic answer ("yes" for the example claim) was awarded a point (equal number of items are positively and negatively keyed). The test was scored for the total number of empathic answers, yielding a range of 0–22. See Table 2 for descriptive statistics.

Working memory capacity. Children's WMC was assessed with the Digit Span subtest of WISC-IV (Wechsler, 2010). The Digit Span test consists of two different tasks. First, in Digits Forward, the child listens to and repeats a sequence of digits spoken aloud by the experimenter. Second, in Digits Backward, the child listens to a digit sequence and repeats it in reverse order. In both parts, the length of each digit sequence increases as the child responds correctly. The test was scored for the total number of correctly recalled digits (range 0–32 points). See Table 2 for descriptive statistics.

2.5. Procedure

Both experiments used the same procedure. All children were tested individually. Upon arrival, they were informed that the experiment assessed reading, and the specific nature of the experiment was explained to them after the experiment. Prior to the reading task, the eye-tracking system was introduced, and the procedure was explained. Children were

Table 2. Descriptive statistics for individual difference and control measures and their intercorrelations

			Correlations				
	M	*SD*	Age	Gender	Empathy index	Digit span	Lukilasse
Age	124.15	3.20					
Gender	0.48	0.50	−.14 [−.17, −.12]				
Empathy index	13.66	2.95	.16 [.14, .19]	.35 [.33, .37]			
Digit span	12.84	2.21	.11 [.08, .13]	.13 [.11, .15]	.09 [.07, .11]		
Lukilasse	82.63	12.90	.15 [.13, .18]	−.17 [−.19, −.14]	−.16 [−.18, −.14]	.20 [.18, .23]	
Maze task	37.68	9.40	.09 [.06, .11]	.13 [.11, .16]	.30 [.28, .32]	.38 [.36, .40]	.29 [.27, .32]

Note: Age is in months. Mean gender is reported as the proportion of girls (n = 97, 46 girls and 51 boys). Values in square brackets indicate the 95% confidence interval.

instructed to read each story for comprehension at their own pace. Each story was presented on a screen, and children pressed the space bar on the keyboard when they finished reading. After each story, two questions were presented, one at a time. Children answered each question by pressing the "Yes" or "No" key on the keyboard. After answering the second question, the next story was presented. The presentation order of the stories was randomised. The reading task was followed by the Digit Span, Lukilasse, and the Index of Empathy for Children and Adolescents. The experimental sessions lasted approximately 30–50 minutes. The reading comprehension task (i.e. the maze task) was completed by the children in their class group after the experiment.

3. Results

3.1. Analyses

Sentence-level measures (Hyönä et al., 2003) were used for analysing the time-course of reading. Descriptive statistics of the measures, as well as their definitions, are presented in Table 3. Separate models were built for each eye-movement measure for the different text regions (see Table 1) and for inference question accuracy. All the measures were analysed for the target phrase, first-pass reading time was analysed for the spillover region, and the probability to look-back was analysed for the critical context. The data were analysed with linear or generalised linear mixed-effects models (Baayen et al., 2008) using the *lme4*

Table 3. Descriptive statistics of the reading and comprehension measures for both story types

		Literal		Irony	
Region	*Measure*	*M*	*SD*	*M*	*SD*
Target phrase					
	First-pass reading time	1840	1127	1873	1163
	Forward fixation time	1410	662	1394	647
	Number of first-pass rereading fixations	1.89	2.61	2.14	2.94
	Probability to look-back	0.19	0.39	0.22	0.41
	Probability to look-from	0.14	0.35	0.16	0.36
Spillover region					
	First-pass reading time	942	667	944	648
Critical context					
	Probability to look-back	0.22	0.42	0.22	0.41
Questions					
	Correct inference	0.90	0.30	0.52	0.50
	Text memory	0.94	0.23	0.94	0.23

Note: First-pass reading time is the summed duration (ms) of fixations made within the sentence during first reading of the sentence. *Forward fixation time* is the summed duration (ms) of fixations landing on unread parts of the sentence during first-pass reading. *Number of first-pass rereading fixations* is the sum of fixations made reinspecting a sentence before moving on. Probability to *look-back* is the proportion of fixations returning to the sentence after the first-pass reading. Probability to *look-from* is the proportion of look-back fixations that were initiated from the sentence. Inference and text memory question accuracies reported are proportions of correct answers.

https://doi.org/10.1017/S0305000926100543 Published online by Cambridge University Press

package (version 1.1–37; Bates et al., 2015) in R statistical software (version 4.4.2; R Core Team, 2024). The reading time measures were skewed and consequently log-transformed by selecting the best fitting transformation using the Box-Cox Power transform (Box & Cox, 1964). Due to the large number of zero values (31%), the number of first-pass rereading fixations was calculated using *glmmTMB* package (Version 1.1.10.; Brooks et al., 2017), as was the case for probability measures with >50% of zero values (i.e. look-back to and look-from the target phrase, and look-back to critical context). The models were tested for zero inflation using *testZeroinflation* function from *DHARMa* package (version 0.4.7.; Hartig, 2024). None of the models showed issues with zero inflation (all p's > .463), and consequently, the zero-inflation part was not fitted.

Story Type (literal versus irony), *Trial Order* (i.e. the order in which the items were presented), and individual difference measures (i.e. *Empathy, WMC*, and *Gender*) and their interactions were included in the models as fixed effects. Technical reading skill, reading comprehension, and relative age were controlled for in the models. In the models, *Story Type* and *Gender* were deviation-coded (see Schad et al., 2020). *Trial Order, Empathy*, *WMC*, and *Age* were fitted in the models as continuous centred fixed effects variables. Only those individual difference measures with a significant interaction with *Story Type* in the model for inference question accuracy were included in the models for the reading measures. To assess whether the correct comprehension of irony is reflected in children's reading times, *Comprehension* (i.e. whether the meaning of the target phrase was understood correctly) was also included in the models as a treatment coded variable, with correct response as the baseline (as in Olkoniemi et al., 2023). Thus, the model intercepts reflect reading where the intended meaning of the target phrase was correctly understood. All the possible three-way interactions including *Story Type* were included in the models for reading. Last, maximal random effects structure was fitted (see Barr et al., 2013). If the model failed to converge, the random structure was trimmed top-down, starting with the correlations between the factors, and continuing by removing random effects with the lowest variance until the model converged.

For the sake of brevity, only significant effects (t and $z > |1.96|$) related to *Story Type* are reported. Significant main effects are reported in the text. Interactions and their 95% CIs are reported in the figures. All final models are reported in Supplementary Tables S1-S8. The data and analysis scripts are available via OSF at https://osf.io/dq8ku/

3.2. Inference and text memory questions

Children's accuracy on the *text memory questions* was close to ceiling and did not differ between literal and ironic stories, $t(96) = -0.16$, $p = .875$, $d = 0.02$. This suggests that the children were attentive to the task and for both story types.

The model for the *inference question accuracy* (see Supplementary Table S1) showed a main effect of *Story Type.* The intended meaning of ironic phrases was harder to comprehend than that of literal phrases, for which accuracy was at ceiling, $\beta = -3.85$, *95% CI* [−4.65, −3.04], $z = -9.38$. Moreover, the model revealed an interaction between *Story Type* and *Trial Order*, which was qualified by two three-way interactions. First, there was an interaction between *Story Type*, *WMC*, and *Trial Order*, $\beta = -0.45$, *95% CI* [−0.66, −0.25], $z = -4.31$ (see Figure 1a), indicating that children with lower WMC showed a steeper increase in inference question accuracy for ironic phrases over the course of the experiment than children with higher WMC. Second, there was an interaction between *Story Type*, *Empathy*, and *Trial Order*, $\beta = 0.36$, *95% CI* [0.15, 0.57], $z = 3.39$ (see

https://doi.org/10.1017/S0305000926100543 Published online by Cambridge University Press

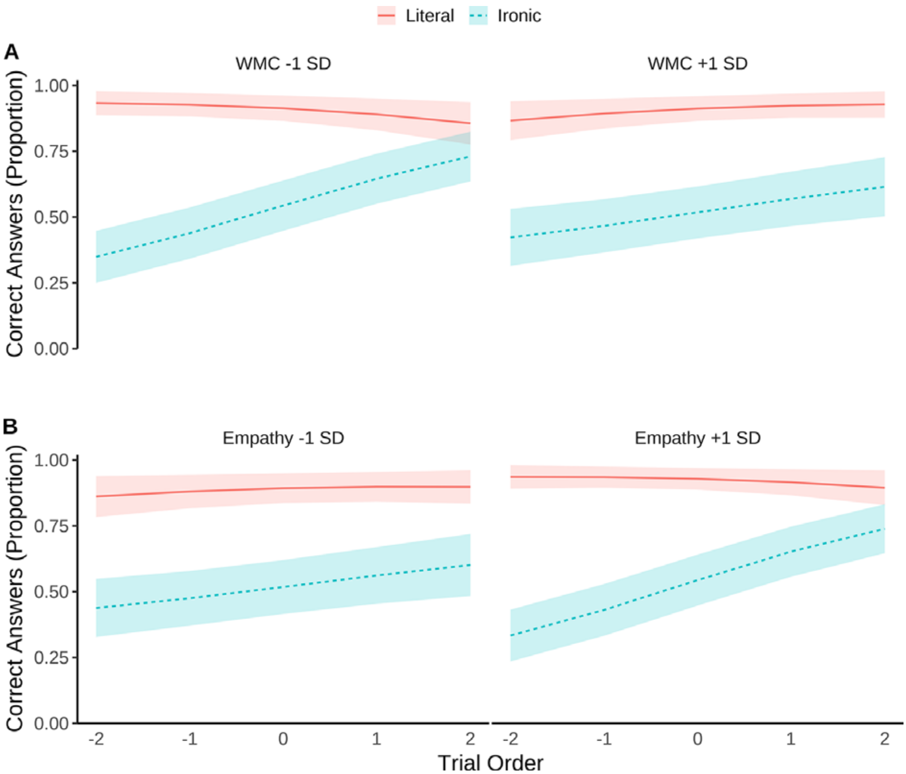

Figure 1. Model estimates for inference question accuracy. *Note:* Panel (a): Model estimates for the interaction between Story Type, WMC, and Trial Order. Panel (b): Model estimates for the interaction between Story Type, Empathy, and Trial Order. In all panels, Trial Order is centred (i.e. value of 0 indicates the middle of the experimental session, negative values indicate the beginning of the experiment, and positive values indicate the end of the experiment). For illustrative purposes, WMC and Empathy scores are divided into high and low groups (± 1 *SD*). The model values are log-back-transformed and shaded areas represent the 95% CI.

Figure 1b). This interaction indicated that children with higher empathy scores showed a steeper increase in inference question accuracy for ironic phrases over the course of the experiment than children with lower empathy scores. There was no significant effect or interaction related to *Gender*, which was therefore not included in further analyses.

3.3. First-pass reading of the target phrase

The model for *first-pass reading time on the target phrase* (see Supplementary Table S2) showed an interaction between *Story Type* and *WMC*, which was qualified by a three-way interaction between *Story Type, Comprehension*, and *WMC*, $\beta = 0.09$, *95% CI* [0.01, 0.16], $t = 2.30$. This interaction indicates that when the intended meaning of the target phrase was correctly interpreted, children with lower WMC showed slower first-pass reading times for the ironic than for the literal target phrase, but this effect was diminished with higher WMC (see Figure 2a). The opposite was true when the intended meaning of the target phrase was interpreted incorrectly: children with higher WMC showed slower first-

https://doi.org/10.1017/S0305000926100543 Published online by Cambridge University Press

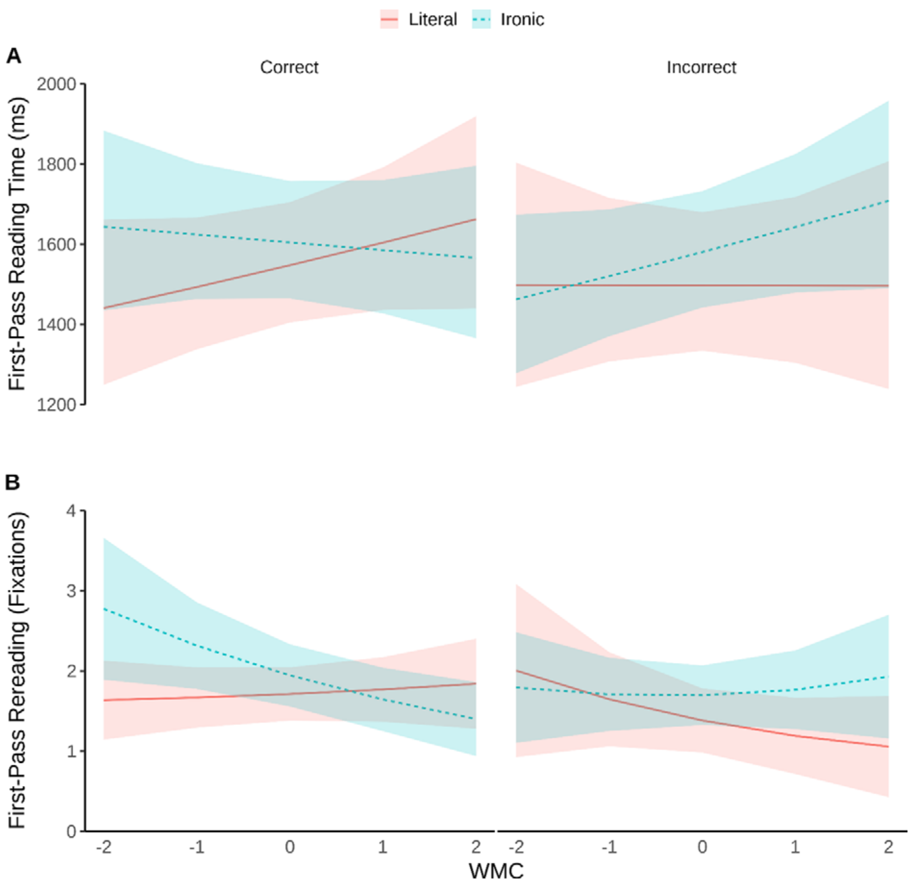

Figure 2. Model estimates for first-pass reading measures on target phrase. *Note:* Panel (a): Model estimates for an interaction between Story Type, WMC (centred), and Comprehension from the model for first-pass reading time on the target phrase. Panel (b): Model estimates for an interaction between Story Type, WMC (centred), and Comprehension from the model for number of first-pass rereading fixations on the target phrase. The model values in panels are log-back-transformed and the shaded areas represent the 95% CI.

pass reading time for the ironic target than for the literal target phrase, whereas children with lower WMC showed virtually no difference between the story types. The model did not reveal any other effects related to *Story Type.*

The model on *forward fixation time on the target phrase* showed no effects related to *Story Type* (see Supplementary Table S3).

The model on *number of first-pass rereading fixations on the target phrase* (see Supplementary Table S4) revealed an interaction between *Story Type* and *WMC*, which was qualified by a three-way interaction between *Story Type, Comprehension*, and *WMC,* $\beta = 0.36$, *95% CI* [0.07, 0.66], $z = 2.41$. The interaction was similar to that observed for first-pass reading time; when the target phrase was correctly interpreted, children with lower WMC reread the ironic target phrase more than the literal target phrase on the first-pass. However, this effect was negligible for children with higher WMC (see Figure 2b). Conversely, when the meaning of the target phrase was interpreted incorrectly, children

https://doi.org/10.1017/S0305000926100543 Published online by Cambridge University Press

with higher WMC reread the ironic target phrase more than the literal one. In contrast, children with lower WMC showed no difference in rereading between the story types. The model did not show any other effects related to *Story Type*.

3.4. Returns to and from the target phrase

The model for the *probability to look-back to the target phrase* (see Supplementary Table S5) showed a main effect of *Story Type*, $\beta = 0.46$, *95% CI* [0.16, 0.76], $z = 3.00$. Children were more likely to look-back to ironic than to literal target phrases. This main effect was qualified by three two-way interactions. First, there was an interaction between *Story Type* and *Comprehension*, $\beta = -0.72$, *95% CI* [−1.42, −0.01], $z = -1.99$, indicating that children were more likely to look-back to the ironic than to the literal target phrase when they had correctly understood the intended meaning of the target phrase (see Figure 3a).

Second, there was an interaction between *Story Type* and *Trial Order*, $\beta = -0.42$, *95% CI* [−0.72, −0.12], $z = -2.79$. The nature of this effect was that at the beginning of the experiment, children were more likely to look-back to ironic than to literal target phrases. Towards the end of the experiment, however, the effect reversed, with children being more likely to look-back to literal target phrases (see Figure 3b).

Last, there was an interaction between *Story Type* and *Empathy*, $\beta = -0.30$, *95% CI* [−0.57, −0.02], $z = -2.12$. This effect indicates that children with lower empathy scores were more likely to look-back to ironic than to literal target phrases, but children with higher empathy scores were more likely to look-back to the literal than to the ironic target phrase (see Figure 3c).

The model on *the probability to look-back from target phrase* showed no effects related to *Story Type* (see Supplementary Table A6).

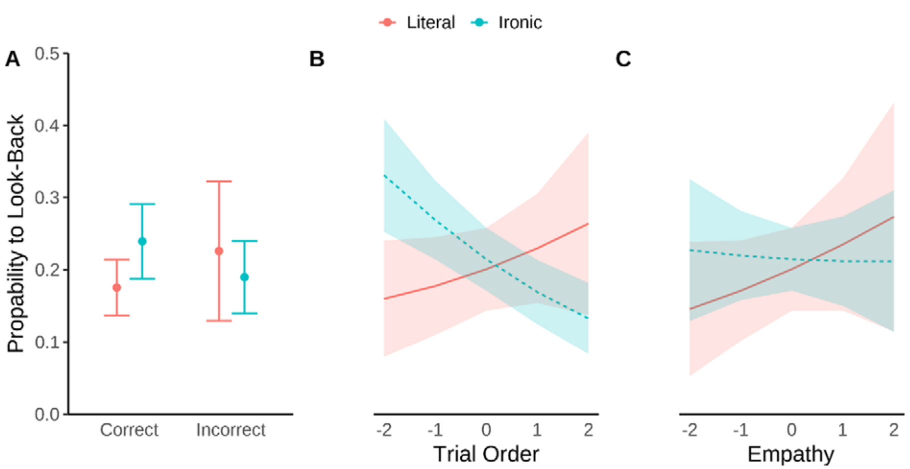

Figure 3. Model estimates for the probability to look-back to target phrase. *Note:* Panel (a): Model estimates for interaction between Story Type and Comprehension. Panel (b): Model estimates for interaction between Story Type and Trial Order (centred). Panel (c): Model estimates for interaction between Story Type and Empathy (centred). The model values in both panels are log-back-transformed; error bars in panel (a), and the shaded areas in panels (b) and (c) represent the 95% CI.

3.5. Processing of spillover region and critical context

The model for *first-pass reading of the spillover region* (see Supplementary Table S7) revealed an interaction between *Story Type, WMC*, and *Comprehension, β* = 0.16, *95% CI* [0.07, 0.25], t = 3.45. The interaction indicates that when the target phrase was correctly interpreted, children with lower WMC had slightly longer first-pass reading times in the spillover region following the ironic target phrase than the literal target phrase, but this effect was negligible with higher WMC (see Figure 4a). However, when the intended meaning of the target phrase was understood incorrectly, children with lower WMC showed lower first-pass reading time on the ironic than on the literal target phrase, and the effect inverted with higher WMC. The model did not show any other effects related to *Story Type*.

The model for *probability to look-back to critical context* revealed an interaction between *Story Type* and *Trial Order*, β = −0.44, *95% CI* [−0.73, −0.15], z = −2.97. The interaction indicates that in the beginning of the experiment, children were more likely to look back to critical context of ironic than literal stories, but the effect turned the other way around towards the end of the experiment (see Figure 4b). The model did not show any other effects related to *Story Type*.

4. Discussion

The aim of the present study was to investigate how individual differences in 10-year-old children's empathy, WMC, and gender are related to their processing and comprehension of written irony. We expected that, in general, children would be slower and less accurate when processing irony than literal language and expected that these effects would be modulated by individual differences. Since the overall processing effects have already been reported in the two previous papers (Olkoniemi et al., 2023, 2025), we only briefly

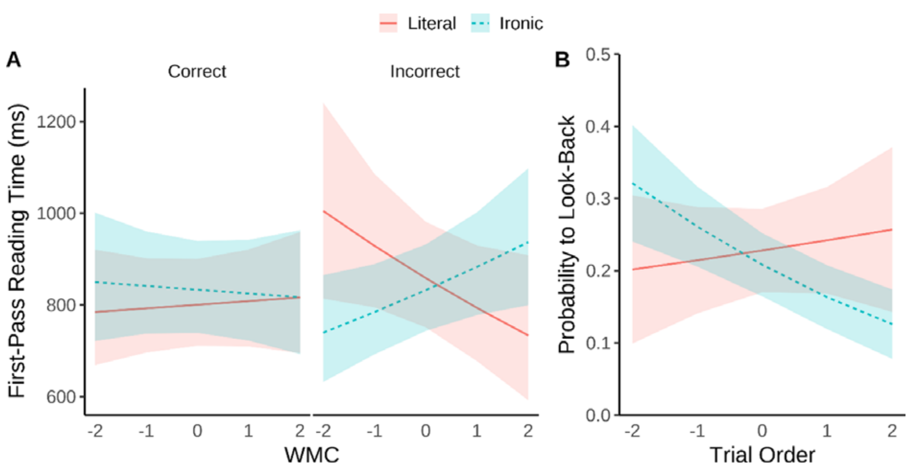

Figure 4. Model estimates for the first-pass reading of the spillover region and probability to look-back to critical context. *Note:* Panel (a): Model estimates for interaction between Story Type, WMC (centred), and Comprehension. Panel (b): Model estimates for interaction between Story Type and Trial Order (centred). The model values in both panels are log-back-transformed, and the shaded areas represent the 95% CI.

summarise them here. The focus of the discussion will be on the individual differences results, presented after the overall results.

Children's processing results confirmed the hypotheses, aligning with previous studies on children and adults (e.g. Barich et al., 2025; Olkoniemi & Kaakinen, 2021). As in the previous studies, increased processing was related to rereading of the ironic phrase (e.g. Barich et al., 2025; Olkoniemi et al., 2023). Successful irony comprehension was particularly related to increased look-backs to the ironic phrase. This is consistent with theoretical assumptions that integrating an ironic interpretation with the context requires additional processing (e.g. Grice, 1975). Furthermore, it aligns with general findings on the role of later rereading, demonstrating that rereading is important for achieving a comprehensive representation of the text (e.g. Hyönä & Nurminen, 2010).

Additionally, as expected, children showed trial effects for irony in late processing measures (i.e. spillover effect, probability to look-back to the target phrase and critical context). This replicates previous findings for children (Olkoniemi et al., 2023, 2025) and adults (Olkoniemi & Kaakinen, 2021; Spotorno & Noveck, 2014). These results suggest that children require additional processing to integrate ironic phrases with context, but repeated exposure to irony creates a prediction of future irony, which allows children to improve their processing performance.

4.1. Individual differences in comprehending written irony

The results confirmed our hypothesis that children's higher empathy skill is associated with higher irony comprehension accuracy, consistent with previous empirical findings demonstrating this relationship (Agostino et al., 2017; Nicholson et al., 2013). This relationship was not, however, observed immediately at the beginning of the experiment. Instead, children with higher empathy skill were better able to improve their comprehension accuracy over the course of the experiment. It seems that better ability to recognise emotions supports irony comprehension and helps children to make more accurate prediction of the use of irony in the future. In addition, children with higher empathy skill were faster at processing ironic phrases throughout the experiment as indicated by a lower likelihood of initiating a look-back to the ironic target phrase. This aligns with eye-tracking studies with adults showing that a more advanced ability to name and recognise emotions is related to faster processing of ironic texts (see Olkoniemi & Kaakinen, 2021, for a review). Our findings demonstrate that, like adults, emotional perspective-taking skill helps children to recognise speakers' emotions in irony, which facilitates processing. However, in our study, higher empathy skill had a relatively small effect on processing of irony per se, which was seen as a slight decrease in the likelihood of looking back to ironic target phrases. Instead, an increase in empathy skill seemed to increase the likelihood of later rereading of literal target phrases. This suggests that, among children with higher empathy, speakers' emotions in literal contexts may require additional verification to confirm whether the nonliteral interpretation is appropriate. As the relationship between children's empathy skill and processing of written irony was explored in this study for the first time, future studies are needed to confirm and further investigate the observed effects.

As expected, children's processing and comprehension of irony were also related to their WMC. Similar to the effect of empathy skill, the effect of WMC on comprehension was not observed at the beginning of the experiment. However, unexpectedly, children with low WMC, and not high WMC, showed improvement in their irony comprehension

https://doi.org/10.1017/S0305000926100543 Published online by Cambridge University Press

over the course of the experiment. High-WMC children remained at the chance level throughout the experiment, showing only a slight increase in comprehension accuracy. This opposes previous findings with children (Godbee & Porter, 2013). Moreover, our processing time results revealed that low-WMC readers showed slower first-pass reading times of an ironic phrase and spillover region when it was correctly comprehended. The finding aligns with theoretical assumptions and experimental findings expecting longer reading times for ironic than for literal phrases. The effect of WMC that was observed for first-pass reading time was driven by first-pass rereading. Regressions made during first-pass reading have generally been shown to support reading comprehension (e.g. Schotter et al., 2014). Moreover, the spillover effect is thought to indicate that the processing of the previous text region was not fully completed when the reader moved on (e.g. Findelsberger et al., 2019). The present result is different from eye-tracking results for adults, in which high-WMC readers show increase in first-pass rereading of the target phrase, and low-WMC readers show increased later rereading (e.g. Olkoniemi & Kaakinen, 2021). The difference is understandable as children generally do more first-pass rereading and less later rereading than adults (Kaakinen et al., 2015). Nevertheless, although our results suggest that the processing strategy of the children with low WMC led to overall better performance, they were likely unaware of their incorrect interpretations. When they misinterpreted the intended meaning of irony, their reading times for ironic and literal items were similar. High-WMC readers showed an opposite pattern, suggesting that they had some awareness that a literal interpretation was not fitting to the context.

There are at least two possible explanations for why we did not observe the hypothesised higher comprehension accuracy for high-WMC readers. First, high-WMC adults have been found to sometimes rely on overly complex resolving strategies, which hampers their performance (Fischer & Holt, 2017; Wiley & Jarosz, 2012). In the context of irony and children with high WMC, this might result in uncertainty about what the correct interpretation is, eventually leading to comprehension failure. Second, it is possible that children with high WMC were able to consider more interpretive possibilities. At 10 years of age, most children still lack extensive knowledge of irony (Olkoniemi et al., 2025). As such, and in line with the parallel-constraint satisfaction model (Pexman, 2008), the interpretive options children activated (e.g. irony, literal statement, or white lie) might have all been equally likely. This might have made their reading slower and also affected their comprehension accuracy during the course of the experiment. This effect has not been observed in previous studies on adults (Olkoniemi & Kaakinen, 2021), possibly because adults have more effective mechanisms for resolving interpretive ambiguity. Future studies are also needed to further explore the role of WMC in children's irony processing.

We found no relationship of child gender to irony comprehension. This is in line with the majority of previous studies that have tested for gender differences in irony comprehension (Garfinkel et al., 2024; Harris & Pexman, 2003; Recchia et al., 2010). This suggests that although previous studies have shown that there can be gender differences in many underlying skills needed for irony comprehension, those differences are not sufficiently strong to influence irony comprehension per se.

5. Conclusions

This study provides the first evidence of the individual differences that are important to children's processing and comprehension of written irony. We showed that both

https://doi.org/10.1017/S0305000926100543 Published online by Cambridge University Press

processing and comprehension were modulated by children's working memory and empathy. Working memory (e.g. Peverill et al., 2016) and empathy (e.g. Ferguson et al., 2024) are skills that continue to develop into late adolescence. Our findings suggest that these factors are important to irony development and thus lead to the expectation that irony comprehension will continue to improve along with the improvement of these related abilities. Of course, our findings are correlational, and additional research, perhaps involving training paradigms (Lee et al., 2021; Olkoniemi et al., 2025), would be needed to help establish causation.

The results of the present study support theories of irony comprehension that consider individual differences (the parallel constraint-satisfaction framework, Pexman, 2008; the predictive coding theory of irony, Fabry, 2021). A useful addition to theoretical models of irony processing would be a detailed developmental component that would generate testable hypotheses about the acquisition of irony and how individual differences affect this process. Future studies should refine and test these theories to determine how perspective-taking and cognitive abilities are related to children's irony processing and thus explain how children overcome the challenge of irony.

Supplementary material. The supplementary material for this article can be found at http://doi.org/10.1017/S0305000926100543.

Acknowledgements. This study was supported by Research Council of Finland grant #338712 awarded to Henri Olkoniemi. Portions of the data were presented at the 22nd European Conference on Eye Movements in Maynooth, Ireland, and at the 24th Conference of the European Society of Cognitive Psychology in Sheffield, UK.

Competing interests. The authors declare none.

Statement of ethical approval. The data were drawn from two previously conducted studies (Olkoniemi et al., 2023, 2025). Ethical approval for data collection in those original studies was granted by the Ethics Committee for Human Sciences at the University of Turku, Finland.

References

Ackerman, B. P. (1983). Form and function in children's understanding of ironic utterances. *Journal of Experimental Child Psychology*, **35**, 487–508. https://doi.org/10.1016/0022-0965(83)90023-1

Agostino, A., **Im-Bolter, N.**, **Stefanatos, A. K.**, & **Dennis, M.** (2017). Understanding ironic criticism and empathic praise: The role of emotive communication. *British Journal of Developmental Psychology*, **35**, 186–201. https://doi.org/10.1111/bjdp.12155

Antoniou, K., & **Milaki, E.** (2021). Irony comprehension in bidialectal speakers. *Modern Language Journal*, **105**, 697–719. https://doi.org/10.1111/modl.12724

Attardo, S. (2000). Irony as relevant inappropriateness. *Journal of Pragmatics*, **32**, 793–826. https://doi.org/10.1016/S0378-2166(99)00070-3

Baayen, R. H., **Davidson, D. J.**, & **Bates, D. M.** (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, **59**, 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Bancroft, T., & **Cook, B.** (Directors) (1998). *Mulan [film]*. United States: The Walt Disney Company.

Barich, T., **Kyriaki, L.**, **Forndran, A.**, **Williamson, P.**, & **Arciuli, J.** (2025). Children's processing of irony during reading in English: An eye-tracking study. *Journal of Experimental Child Psychology*, **254**, 106202. https://doi.org/10.1016/j.jecp.2025.106202

Barr, D. J., **Levy, R.**, **Scheepers, C.**, & **Tily, H. J.** (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, **68**, 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., **Mächler, M.**, **Bolker, B.**, & **Walker, S.** (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, **67**, 1–48. https://doi.org/10.18637/jss.v067.i01

Blythe, H. I., & **Joseph, H. S. S. L.** (2011). Children's eye movements during reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 643–662). Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199539789.013.0036

Booth, W. C. (1974*). A rhetoric of irony*. Chicago: University of Chicago Press.

Box, G. E. P., & **Cox, D. R.** (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, **26**, 211–252. https://doi.org/10.1111/j.2517-6161.1964.tb00553.x

Brooks, M. E., **Kristensen, K.**, **van Benthem, K. J.**, **Magnusson, A.**, **Berg, C. W.**, **Nielsen, A.**, **Skaug, H. J.**, **Mächler, M.**, & **Bolker, B. M.** (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal*, **9**, 378–400. https://doi.org/10.32614/RJ-2017-066

Bryant, B. K. (1982). An index of empathy for children and adolescents. *Child Development*, **53**, 413–425. https://doi.org/10.2307/1128984

Caillies, S., **Bertot, V.**, **Motte, J.**, **Raynaud, C.**, & **Abely, M.** (2014). Social cognition in ADHD: Irony understanding and recursive theory of mind. *Research in Developmental Disabilities*, **35**, 3191–3198. https://doi.org/10.1016/j.ridd.2014.08.002

Capelli, C. A., **Nakagawa, N.**, & **Madden, C. M.** (1990). How children understand sarcasm: The role of context and intonation. *Child Development*, **61**, 1824–1841. https://doi.org/10.2307/1130840

Climie, E. A., & **Pexman, P. M.** (2008). Eye gaze provides a window on children's understanding of verbal irony. *Journal of Cognition and Development*, **9**, 257–285. https://doi.org/10.1080/15248370802247939

Creusere, M. A. (1999). Theories of adults' understanding and use of irony and sarcasm: Applications to and evidence from research with children. *Developmental Review*, **19**, 213–262. https://doi.org/10.1006/drev.1998.0474

Demorest, A., **Myer, C.**, **Phelps, E.**, **Gardner, H.**, & **Winner, E.** (1984). Words speak louder than actions: Understanding deliberately false remarks. *Child Development*, **55**, 1527–1534. https://doi.org/10.2307/1130022

Dews, S., **Winner, E.**, **Kaplan, J.**, **Rosenblatt, E.**, **Hunt, M.**, **Lim, K.**, **McGovern, A.**, **Qualter, A.**, & **Smarsh, B.** (1996). Children's understanding of the meaning and functions of verbal irony. *Child Development*, **67**, 3071–3085. https://doi.org/10.1111/j.1467-8624.1996.tb01903.x

Fabry, R. E. (2021). Getting it: A predictive processing approach to irony comprehension. *Synthese*, **198**, 6455–6489. https://doi.org/10.1007/s11229-019-02470-9

Ferguson, H. K., **De Lillo, M.**, **Woodrow-Hill, C.**, **Foley, R.**, & **Bradford, E. E. F.** (2024). Neural empathy mechanisms are shared for physical and social pain, and increase from adolescence to older adulthood. *Social Cognitive and Affective Neuroscience*, **19**, nsae080. https://doi.org/10.1093/scan/nsae080

Filik, R., & **Moxey, L. M.** (2010). The on-line processing of written irony. *Cognition*, **116**, 421–436. https://doi.org/10.1016/j.cognition.2010.06.005

Filippova, E., & **Astington, J. W.** (2010). Children's understanding of social-cognitive and social-communicative aspects of discourse irony. *Child Development*, **81**, 913–928. https://doi.org/10.1111/j.1467-8624.2010.01442.x

Findelsberger, E., **Hutzler, F.**, & **Hawelka, S.** (2019). Spill the load: Mixed evidence for a foveal load effect, reliable evidence for a spillover effect in eye-movement control during reading. *Attention, Perception, and Psychophysics*, **81**, 1442–1453. https://doi.org/10.3758/s13414-019-01689-5

Fischer, H., & **Holt, D. V.** (2017). When high working memory capacity is and is not beneficial for predicting nonlinear processes. *Memory and Cognition*, **45**, 404–412. https://doi.org/10.3758/s13421-016-0665-0

Friston, K., & **Frith, C.** (2015). A duet for one. *Consciousness and Cognition*, **36**, 390–405. https://doi.org/10.1016/j.concog.2014.12.003

Friston, K., **Kilner, J.**, & **Harrison, L. A.** (2006). A free energy principle for the brain. *Journal of Physiology – Paris*, **100**, 70–87. https://doi.org/10.1016/j.jphysparis.2006.10.001

Fuchs, J. (2023). 40 years of research into children's irony comprehension: A review. *Pragmatics & Cognition*, **30**, 1–30. https://doi.org/10.1075/pc.22015.fuc

Garfinkel, S., **Rowe, M. L.**, **Bosacki, S.**, & **Banasik-Jemielniak, N.** (2024). Mom said it in quotation marks!" irony comprehension and metapragmatic awareness in 8-year-olds. *Journal of Child Language*, **51**, 485–508. https://doi.org/10.1017/S0305000923000399

Gibbs, R. W., & **Colston, H. L.** (2007). The future of irony studies. In R. W. Gibbs & H. L. Colston (Eds.), *Irony in language and thought: A cognitive science reader* (pp. 581–593). New York: Lawrence Erlbaum Associates.

Giora, R. (1999). On the priority of salient meanings: Studies of literal and figurative language. *Journal of Pragmatics*, **31**, 919–929. https://doi.org/10.1016/S0378-2166(98)00100-3

Glenwright, M., **Tapley, B.**, **Rano, J. K. S.**, & **Pexman, P. M.** (2017). Developing appreciation for sarcasm and sarcastic gossip: It depends on perspective. *Journal of Speech, Language, and Hearing Research*, **60**, 3295–3309. https://doi.org/10.1044/2017_JSLHR-L-17-0058

Godbee, K., & **Porter, M.** (2013). Comprehension of sarcasm, metaphor and simile in Williams syndrome. *International Journal of Language & Communication Disorders*, **48**, 651–665. https://doi.org/10.1111/1460-6984.12037

Grice, H. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics 3: Speech acts* (pp. 41–58). New York: Academic Press.

Hajovsky, D. B., **Caemmerer, J. M.**, & **Mason, B. A.** (2021). Gender differences in children's social skills growth trajectories. *Applied Developmental Science*, **26**, 488–503. https://doi.org/10.1080/10888691.2021.1890592

Harris, M., & **Pexman, P. M.** (2003). Children's perceptions of the social functions of verbal irony. *Discourse Processes*, **33**, 147–165. https://doi.org/10.1207/S15326950DP3603_1

Hartig, F. (2024). *DHARMa: Residual diagnostics for hierarchical (multi-level/mixed) regression models*. https://cran.r-project.org/package=DHARMa

Hyönä, J., **Lorch, R. F.**, & **Rinck, M.** (2003). Eye movement measures to study global text processing. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 313–334). Amsterdam: Elsevier. https://doi.org/10.1016/B978-044451020-4/50018-9

Hyönä, J., & **Nurminen, A.-M.** (2010). Do adult readers know how they read? Evidence from eye movement patterns and verbal reports. *British Journal of Psychology*, **97**, 31–50. https://doi.org/10.1348/000712605X53678

Häyrinen, T., **Serenius-Sirve, S.**, & **Korkman, M.** (2013). *Lukilasse 2: Lukemisen, kirjoittamisen ja laskemisen seulontatesti 1.–6. Vuosiluokille. [Lukilasse 2: Screening test for reading, writing and calculus for 1st to 6th grades.]*. Helsinki: Hogrefe Psykologien Kustannus Oy.

Kaakinen, J. K., **Lehtola, A.**, & **Paattilammi, S.** (2015). The influence of a reading task on children's eye movements during reading. *Journal of Cognitive Psychology*, **27**, 640–656. https://doi.org/10.1080/20445911.2015.1005623

Kaakinen, J. K., **Olkoniemi, H.**, **Kinnari, T.**, & **Hyönä, J.** (2014). Processing of written irony: An eye movement study. *Discourse Processes*, **51**, 287–311. https://doi.org/10.1080/0163853X.2013.870024

Köder, F., & **Falkum, I. L.** (2021). Irony and perspective-taking in children: The roles of norm violations and tone of voice. *Frontiers in Psychology*, **2**, 624604. https://doi.org/10.3389/fpsyg.2021.624604

Lee, K., **Sidhu, D. M.**, & **Pexman, P. M.** (2021). Teaching sarcasm: Evaluating metapragmatic training for typically-developing children. *Canadian Journal of Experimental Psychology*, **75**, 139–145. https://doi.org/10.1037/cep0000228

Loukusa, S., & **Leinonen, E.** (2008). Development of comprehension of ironic utterances in 3- to 9-year-old Finnish-speaking children. *Psychology of Language and Communication*, **12**, 55–69. https://doi.org/10.2478/v10057-008-0003-0

Lyu, A., **Abel, L.**, & **Cheong, A. M. Y.** (2023). Effect of habitual reading direction on saccadic eye movements: A pilot study. *PLoS One*, **18**, e0286801. https://doi.org/10.1371/journal.pone.0286801

Manu, M., **Torppa, M.**, **Vasalampi, K.**, **Lerkkanen, M.**, **Poikkeus, A.**, & **Niemi, P.** (2023). Reading development from kindergarten to age 18: The role of gender and parental education. *Reading Research Quarterly*, **58**, 505–538. https://doi.org/10.1002/rrq.518

McPhillips, M., & **Jordan-Black, J.-A.** (2009). The effect of month of birth on the attainments of primary and secondary school pupils. *British Journal of Educational Psychology*, **79**, 419–438. https://doi.org/10.1348/978185408X380199

Nicholson, A., **Whalen, J. M.**, & **Pexman, P. M.** (2013). Children's processing of emotion in ironic language. *Frontiers in Psychology*, **4**, 691. https://doi.org/10.3389/fpsyg.2013.00691

Olkoniemi, H., **Halonen, S.**, **Pexman, P. M.**, & **Häikiö, T.** (2023). Children's processing of written irony: An eye movement study. *Cognition*, **238**, 105508. https://doi.org/10.1016/j.cognition.2023.105508

Olkoniemi, H., **Häikiö, T.**, **Merinen, M.**, **Manninen, J.**, **Laine, M.**, & **Pexman, P. M.** (2025). Learning irony in school: Effects of metapragmatic training. *Journal of Child Language*. Advance online publication. https://doi.org/10.1017/S0305000925000054

Olkoniemi, H., **Johander, E.**, & **Kaakinen, J. K.** (2019a). The role of look-backs in the processing of written sarcasm. *Memory & Cognition*, **47**, 87–105. https://doi.org/10.3758/s13421-018-0852-2

Olkoniemi, H., & **Kaakinen, J. K.** (2021). Processing of irony in text: A systematic review of eye-tracking studies. *Canadian Journal of Experimental Psychology*, **75**, 99–106. https://doi.org/10.1037/cep0000216

Olkoniemi, H., **Mézière, D. C.**, & **Kaakinen, J. K.** (2024). Comprehending irony in text: Evidence from scanpaths. *Discourse Processes*, **1-2**, 6–20. https://doi.org/10.1080/0163853X.2023.2266964

Olkoniemi, H., **Ranta, H.**, & **Kaakinen, J. K.** (2016). Individual differences in the processing of written sarcasm and metaphor: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **42**, 433–450. https://doi.org/10.1037/xlm0000176

Olkoniemi, H., **Strömberg, V.**, & **Kaakinen, J. K.** (2019b). The ability to recognise emotions predicts the time-course of sarcasm processing: Evidence from eye movements. *Quarterly Journal of Experimental Psychology*, **72**, 1212–1223. https://doi.org/10.1177/1747021818807864

Parola, A., & **Bosco, F. M.** (2022). An eye-tracking investigation of the cognitive processes involved in the comprehension of simple and complex communicative acts. *Quarterly Journal of Experimental Psychology*, **75**, 1976–1995. https://doi.org/10.1177/17470218221079629

Perner, J., & **Wimmer, H.** (1985). "John *thinks* that Mary *thinks* that…": Attribution of second-order beliefs by 5- to 10-year-old children. *Journal of Experimental Child Psychology*, **39**, 437–471. https://doi.org/10.1016/0022-0965(85)90051-7

Peverill, M., **McLaughlin, K. A.**, **Finn, A. S.**, & **Sheridan, M. A.** (2016). Working memory filtering continues to develop into late adolescence. *Developmental Cognitive Neuroscience*, **18**, 78–88. https://doi.org/10.1016/j.dcn.2016.02.004

Pexman, P. M. (2008). It's fascinating research: The cognition of verbal irony. *Current Directions in Psychological Science*, **17**, 286–290. https://doi.org/10.1111/j.1467-8721.2008.00591.x

Pexman, P. M., & **Glenwright, M.** (2007). How do typically developing children grasp the meaning of verbal irony? *Journal of Neurolinguistics*, **20**, 178–196. https://doi.org/10.1016/j.jneuroling.2006.06.001

Pexman, P. M., **Glenwright, M.**, **Hala, S.**, **Kowbel, S. L.**, & **Jungen, S.** (2006). Children's use of trait information in understanding verbal irony. *Metaphor and Symbol*, **21**, 39–60. https://doi.org/10.1207/s15327868ms2101_3

Proverbio, A. M. (2023). Sex differences in the social brain and in social cognition. *Journal of Neuroscience Research*, **101**, 730–738. https://doi.org/10.1002/jnr.24787

R Core Team (2024). R: A language and environment for statistical computing. Retrieved from: http://www.R-project.org/

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, **62**, 1457–1506. https://doi.org/10.1080/17470210902816461

Recchia, H. E., **Howe, N.**, **Ross, H. S.**, & **Alexander, S.** (2010). Children's understanding and production of verbal irony in family conversations. *British Journal of Developmental Psychology*, **28**, 255–274. https://doi.org/10.1348/026151008X401903

Riess, H. (2017). The science of empathy. *Journal of Patient Experience*, **4**, 74–77. https://doir.org/10.1177/2374373517699267

Ronimus, M., **Tolvanen, A.**, & **Hautala, J.** (2022). The roles of motivation and engagement in computer-based assessment of children's reading comprehension. *Learning and Individual Differences*, **98**, 102197. https://doi.org/10.1016/j.lindif.2022.102197

Rothermich, K., **Caivano, O.**, **Knoll, L. J.**, & **Talwar, V.** (2020). Do they really mean it? Children's inference of speaker intentions and the role of age and gender. *Language and Speech*, **63**, 689–712. https://doi.org/10.1177/0023830919878742

Rothermich, K., & **Pell, M. D.** (2015). Introducing RISC: A new video inventory for testing social perception. *PLoS One*, **10**, e0133902. https://doi.org/10.1371/journal.pone.0133902

Schad, D. J., **Vasishth, S.**, **Hohenstein, S.**, & **Kliegl, R.** (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language*, **110**, 104038. https://doi.org/10.1016/j.jml.2019.104038

Schotter, E. R., **Tran, R.**, & **Rayner, K.** (2014). Don't believe what you read (only once): Comprehension is supported by regressions during reading. *Psychological Science*, **25**, 1218–1226. https://doi.org/10.1177/0956797614531148

Spotorno, N., & **Noveck, I. A.** (2014). When is irony effortful? *Journal of Experimental Psychology: General*, **143**, 1649–1665. https://doi.org/10.1037/a0036630

SR Research (2025, Oct 27). EyeLink eye trackers comparison chart. https://www.sr-research.com/eyelink-eye-trackers-comparison-chart/

Tiv, M., **Deodato, F.**, **Rouillard, V.**, **Wiebe, S.**, & **Titone, D.** (2020). Second language experience impacts first language irony comprehension among bilingual adults. *Canadian Journal of Experimental Psychology*, **75**, 126–132. https://doi.org/10.1037/cep0000230

Wechsler, D. (2010). *WISC IV - Wechsler intelligence scale for children* (4th ed.). Finnish version. Helsinki: Psykologien Kustannus Oy.

Wiley, J., & **Jarosz, A. F.** (2012). Working memory capacity, attentional focus, and problem solving. *Current Directions in Psychological Science*, **21**, 258–262. https://doi.org/10.1177/0963721412447622

Wilson, D. (2013). Irony comprehension: A developmental perspective. *Journal of Pragmatics*, **59**, 40–56. https://doi.org/10.1016/j.pragma.2012.09.016

Winner, E., & **Gardner, H.** (1993). Metaphor and irony: Two levels of understanding. In A. Ortony (Ed.), *Metaphor and thought* (pp. 425–443). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139
173865.021

Winner, E., & **Leekam, S.** (1991). Distinguishing irony from deception: Understanding the speaker's second-order intention. *British Journal of Developmental Psychology*, **9**, 257–270. https://doi.org/10.1111/j.2044-835X.1991.tb00875.x

Zajączkowska, M., & **Abbot-Smith, K.** (2020). Sure I'll help—I've just been sitting around doing nothing at school all day": Cognitive flexibility and child irony interpretation. *Journal of Experimental Child Psychology*, **199**, 104942. https://doi.org/10.1016/j.jecp.2020.104942

Zald, D. H., & **Andreotti, C.** (2010). Neuropsychological assessment of the orbital and ventromedial prefrontal cortex. *Neuropsychologia*, **48**, 3377–3391. https://doi.org/10.1016/j.neuropsychologia.2010.08.012

Cite this article: Olkoniemi, H., Häikiö, T., Laine, M., & Pexman, P.M. (2026). Who gets it? Explaining variability in children's written irony comprehension. *Journal of Child Language* 1–24, https://doi.org/10.1017/S0305000926100543