



HAL
open science

When irony is faster than its literal control: The role of mindreading during irony comprehension

Camilo R Ronderos, John M Tomlinson, Ira Noveck

► To cite this version:

Camilo R Ronderos, John M Tomlinson, Ira Noveck. When irony is faster than its literal control: The role of mindreading during irony comprehension. *Journal of Experimental Psychology: Learning, Memory and Cognition*, In press. hal-03913974

HAL Id: hal-03913974

<https://hal.science/hal-03913974>

Submitted on 27 Dec 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 When irony is faster than its literal control: The role of mindreading during irony
2 comprehension

3 Camilo R. Ronderos^{1,2}, John M. Tomlinson. Jr.³, & Ira Noveck⁴

4 ¹ University of Oslo, Department of Philosophy, Classics, History of Art and Ideas

5 ² Humboldt-Universität zu Berlin, Institut für deutsche Sprache und Linguistik

6 ⁴ Laboratoire de Linguistique Formelle, UMR 7110 -Université de Paris & CNRS

7 ³ Leibniz-ZAS Berlin

8 Author Note

9 The authors declare no competing interests. All data, pre-registration forms and
10 analysis scripts are available on the associated OSF project: <https://osf.io/vgkst/>

11 Correspondence concerning this article should be addressed to Camilo R. Ronderos,
12 Blindernveien 31 Georg Morgenstiernes hus, room 533, Oslo. E-mail: camilorr@uio.no

Abstract

13

14 Irony is a heavily context-dependent pragmatic phenomenon. But what is it about context
15 that facilitates or blocks irony comprehension? Based on the echoic account, we suggest that
16 a context facilitates irony comprehension when it makes manifest a speaker's intentions and
17 attitude, i.e., when a context makes it easy for participants to engage their mindreading
18 abilities. In two pre-registered self-paced reading experiments, we investigated the
19 comprehension of sentences in English that could be understood as ironic or literal,
20 according to the story frame that participants read leading to the target sentence. In
21 Experiment 1, we found that when the story frames prevent participants from anticipating
22 the speaker's intention, literal readings of critical sentences are - not surprisingly - faster
23 than ironic ones. Importantly, when the story frames gave access to the speaker's intentions,
24 we find cases in which ironic readings are actually faster than literal ones, resulting in a
25 novel finding for the irony comprehension literature. Further, when the speaker was
26 described as having a sincere attitude towards their utterance, participants tended to
27 understand the utterances literally. They tended to understand them ironically when it was
28 not clear what the speaker's attitude was. In Experiment 2 we investigated whether the
29 findings of Experiment 1 could be linked to individual differences in participants'
30 mindreading abilities. We found that participants who scored higher on a standard Theory
31 of Mind task (the 'Reading the mind in the Eyes' task) were significantly more likely to
32 derive ironic - but not literal - interpretations. We see these results as supporting the echoic
33 account of irony comprehension. This work discusses the relevance of our findings to the
34 long-standing debate on the processing effort of ironic vs. literal sentences.

35

Keywords: irony comprehension, mindreading, echoic account, Theory of Mind,

36

experimental pragmatics, figurative language

37

Word count: 8944

38 When irony is faster than its literal control: The role of mindreading during irony
39 comprehension

40 **Introduction**

41 It is often the case that people mean something very different from what they actually
42 say. Take the conversation in (1):

43 (1) (A) *Chris: Sorry, could my daughter play this guitar?*

44 (B) *Music store owner: Yes, this guitar is here for everyone to play with.*

45 When taken literally, (1B) could just be a sincere answer to a polar (yes-no) question.
46 However, if the store owner sees that Chris's daughter is a small child and knows that the
47 guitar is incredibly expensive, he might actually mean something quite different. The store
48 owner might wish to convey that the question is ridiculous and by no means can Chris's
49 daughter play the guitar. This would be an instance of **verbal irony**: a language strategy
50 through which an indirect, evaluative utterance is communicated with a proposition that
51 stands in some type of opposition to the speaker's intentions (see Bryant, 2012; Pexman,
52 2008).

53 It is clear that one must go beyond the literal meaning of (1B) to understand it
54 ironically. But what exactly does a comprehender need to do for this to happen? One well
55 known approach to irony interpretation, *the echoic account* (Jorgensen, Miller, & Sperber,
56 1984; Sperber & Wilson, 1981; Wilson & Sperber, 1992, 2012), sees irony as a type of
57 attributive use of language. When using irony, a speaker does not endorse their own
58 utterance, but instead implicitly attributes it to someone else or to some normative
59 expectation. This amounts to expressing a dissociative attitude towards the belief
60 articulated in the utterance. This analysis suggests that to understand irony, the
61 comprehender must ultimately accomplish two things: (i) gain access to the speaker's
62 informative intention (*what* it is that the speaker wants to convey) and; (ii) detect the
63 speaker's attitude towards their own proposition (i.e., to capture *that* the speaker wants their

64 audience to get the informative intention through an ironic attitude). These two features
65 combined illustrate that irony comprehension crucially involves a form of reasoning about
66 mental states that allows a comprehender to interpret a speaker's behavior. Such reasoning
67 generally falls under the umbrella term of mindreading (Nichols & Stich, 2003; Spaulding,
68 2020) or Theory of Mind (Baron-Cohen, Leslie, & Frith, 1985).

69 Prior tests of the echoic account have shown that listeners can more readily process
70 ironic utterances if there is an explicit echo in the discourse context (Gibbs, 1986; Jorgensen,
71 Miller, & Sperber, 1984; Turcan & Filik, 2017). What is not known, however, is how the two
72 previously mentioned types of mindreading skills - considering a speaker's informative
73 intention and the speaker's attitude towards their proposition - affect irony processing. If a
74 context facilitates these types of reasoning, will irony be more readily understood? Besides
75 serving as a test of the echoic account, answering this question puts one in a position to
76 address one of the longest-running debates in the processing literature on irony: Are ironic
77 sentences harder to process than their literal controls?

78 In this work, we propose that mindreading can have a *variable* effect on ironic readings
79 of utterances with respect to literal readings of one and the same sentence. In what follows,
80 we first review the evidence linking mindreading to irony comprehension. We then discuss
81 the psycholinguistic findings that investigate the processing effort of ironic, relative to literal,
82 utterances. Then, we present our two experiments and discuss them in the light of the issues
83 raised in the Introduction.

84 **Irony and mindreading**

85 According to the echoic account, when an addressee successfully understands an ironic
86 utterance, they understand that the speaker is attributing this utterance or thought to
87 someone else while simultaneously expressing a dissociative attitude towards it. This means
88 that irony comprehension should require the ability to generate second-order
89 metarepresentations (e.g., a thought about a thought, as in the sentence *Miguel thinks that*

90 *Luisa is upset that Paula is leaving*). This ability is believed to be an integral part of adult
91 mindreading skills (Allott, 2017; Sperber & Wilson, 2002, i.a.).

92 Previous studies have shown that difficulties in generating second-order
93 metarepresentations (due to either brain lesions or atypical neurological development)
94 correlate with difficulties in understanding irony, but not with difficulties in understanding
95 literal language (F. G. E. Happé, 1993; McDonald, 2000). It has also been shown that brain
96 regions typically associated with mindreading activity display increased activation during an
97 irony comprehension task relative to literal controls (Spotorno, Koun, Prado, Van Der Henst,
98 & Noveck, 2012). These studies are in line with the echoic account's predictions regarding
99 the involvement of mindreading in irony comprehension (for a summary, see Noveck, 2018).
100 However, it is unclear from these studies whether this involvement is binary (you either have
101 it or you don't) or whether mindreading can have a graded effect on comprehension.

102 Spotorno and Noveck (2014) were the first to demonstrate that engaging in
103 mindreading skills during irony comprehension is arguably a matter of degree:
104 Comprehenders can be shown to progressively anticipate mindreading situations over the
105 course of an experiment and, as a result, understand irony more readily (as measured by
106 reduced reading times for ironic utterances) by the end of an experimental session. It is
107 important to note that, though this finding suggests an involvement of mindreading, it differs
108 from what the echoic account would predict in a critical way. While Spotorno and Noveck
109 (2014) showed that comprehenders' processing effort can be reduced through repeated
110 encounters of irony, the echoic account would state that the processing effort of a single ironic
111 sentence will depend on that sentence's communicative context, not on whether different
112 ironic sentences have been encountered before. In other words, the echoic account does not
113 state that people engage in mindreading to anticipate 'irony' as a trope. Comprehenders do
114 so to anticipate the beliefs and intentions of the speaker that lay behind a single ironic
115 utterance. A different type of contextual manipulation is therefore needed to investigate
116 whether mindreading has a variable effect on processing irony in an individual trial.

117 Processing effort of ironic vs. literal utterances

118 If, as we hypothesize, mindreading can have a variable effect on irony comprehension,
119 this should be reflected during online processing. If comprehenders have strong evidence as
120 to the nature of the speaker's intentions and beliefs, it should be easier for them to
121 understand irony compared to comparable cases that do not provide such evidence.
122 Behaviorally, this should reveal that the processing effort of irony could in fact be more
123 efficient than that of literal readings under certain mindreading-related conditions. For
124 example, a sentence understood ironically in a context rich in mindreading-facilitating cues
125 should be more readily understood than a sentence understood literally in a context deprived
126 of evidence pointing to the speaker's intentions and beliefs. In other words, processing effort
127 of ironic (and literal) utterances is constrained by a comprehender's expectations of the
128 speaker's intended meaning (see Degen & Tanenhaus, 2019 for a related argument).

129 Investigating the variable role of mindreading during irony comprehension therefore
130 bears on one of the central topics in irony research: The debate on processing effort of ironic
131 relative to literal language. Broadly speaking, there are two camps in this debate. On one
132 side sits the *contextualist* camp, which states that context influences processing such that an
133 ironic sentence can be understood just as easily as a literal equivalent. This view has its
134 origins in Gibbs (1986), who claimed that understanding irony could happen 'directly'
135 without first deriving a literal interpretation, in opposition to previous accounts (Grice, 1989;
136 Searle, 1979). Gibbs's approach, known as the Direct Access view (Gibbs, 1994, 2002), is
137 complemented by the Constraint Satisfaction view (Pexman, 2008; Pexman, Ferretti, & Katz,
138 2000), which states that multiple factors can influence processing of irony in parallel, often
139 resulting in ironic sentences being understood just as fast as their literal counterparts
140 (Ivanko & Pexman, 2003; Katz, Blasko, & Kazmerski, 2004).

141 On the other side sits the *context-independence* camp, which states that ironic
142 sentences typically require more processing resources than their literal counterparts
143 regardless of context (Giora & Fein, 1999; Giora et al., 2007; Giora, Fein, & Schwartz, 1998;

144 Schwoebel, Dews, Winner, & Srinivas, 2000). A prominent representative of this view is the
145 Graded Salience Hypothesis (Giora, 2003). It states that for any utterance, the most salient
146 meaning will be processed by default. While salience is determined by the utterance's
147 familiarity, stereotypicality, prototypicality, and frequency (among other factors), the most
148 salient interpretation usually coincides with the utterance's literal meaning. If a salient
149 meaning is found to be incompatible with context, a secondary, non-salient meaning is
150 computed. Irony is normally non-salient (but see Giora, Drucker, Fein, & Mendelson, 2015
151 for some exceptions), so it is only understood after deriving the literal meaning, resulting in
152 more processing effort compared to that of understanding a literal utterance (Filik, Howman,
153 Ralph-Nearman, & Giora, 2018; Filik & Moxey, 2010; Giora et al., 2007).

154 Tests of these accounts have usually consisted in looking for contextual cues that may
155 or may not ease comprehension of an ironic relative to a literal utterance. This has resulted
156 in conflicting evidence (e.g. Filik & Moxey, 2010; Ivanko & Pexman, 2003; Katz, Blasko, &
157 Kazmerski, 2004; Schwoebel, Dews, Winner, & Srinivas, 2000). Some elements of context -
158 such as the presence of an 'echoed' antecedent (Ṫurcan & Filik, 2017), explicitly introducing
159 a character as sarcastic (Ṫurcan, Howman, & Filik, 2020), or the association of one character
160 with sarcasm throughout an entire experiment (Regel, Coulson, & Gunter, 2010) - seem to
161 facilitate processing. While others - such as the presence of a previous sarcastic utterance by
162 the speaker (Giora et al., 2007) or explicit mention of the speaker's expectations (Ṫurcan &
163 Filik, 2016) - do not.

164 Given the current state of the debate, there is no unified account whose predictions
165 adequately explain these empirical findings. We offer a different approach. Instead of
166 focusing on the specific features of a context that might speed up processing, we examine the
167 effect of context only in as much as it can help participants in an experiment anticipate the
168 speaker's informative intention as well as the speaker's attitude. In other words, we suggest
169 that examining two ways in which participants are encouraged to engage in mindreading
170 abilities will help us understand how the processing effort linked to irony varies, relative to

171 the processing effort linked to literal readings.

172 **The variable effect of mindreading on irony comprehension**

173 Let us revisit example (1). In the event that we know more about the store owner –
174 e.g., that he has no intention of letting a young girl play with a very expensive guitar – the
175 reader will likely expect the store owner’s answer to be ‘no.’ The interpretation – and
176 processing effort – of (1B) will thus be determined by how strong such expectations are.
177 Knowing the store owner’s attitude when he speaks is also a cue to an ironic reading. A
178 reader who is further told that the speaker has a tendency to speak insincerely (e.g. jokingly)
179 is more likely to read (1B) ironically. In short, the more strongly that a comprehender
180 believes to know the store owner’s intentions and attitude, the easier it should be to
181 interpret (1B) as ironic or not.

182 This leads to the goal of the current study. Based on the predictions of the echoic
183 account, we investigate how processing effort of irony varies as a function of features related
184 to irony and mindreading. Concretely, we test the following two hypotheses. First, we
185 hypothesize that a context facilitates irony comprehension when it provides comprehenders
186 with a deeper understanding of a speaker’s intention as described through a story frame and
187 by giving explicit information about a speaker’s attitude. If this hypothesis is on the right
188 track, one should be able to manipulate such anticipations in such a way that ironic (as well
189 as sincere) readings of identical sentences can be equally facilitated. In the event that
190 mindreading-rich contexts do not facilitate the processing of ironic readings compared to
191 mindreading-poor contexts, it would speak against the echoic account and offer support to
192 views that see irony comprehension as a generally more effortful process than understanding
193 literal sentences, regardless of contextual bias (e.g., Giora et al., 2007; Schwoebel, Dews,
194 Winner, & Srinivas, 2000). Second, we hypothesize that, if a facilitatory effect of context is
195 in fact linked to mindreading, it should be more pronounced for comprehenders who are
196 particularly apt at using their mindreading abilities relative to those who are less so.

197 Alternatively, if there is no connection between individual differences in mindreading and an
198 effect of context, it would suggest that the way in which comprehenders integrate contextual
199 cues with an utterance during irony interpretation does not necessarily require reasoning
200 about a speaker's informative intention and attitude towards their own utterance.

201 Concretely, we first validated a narrative context that can lead to either an ironic or a
202 literal reading of a sentence while allowing for reading times measures and comprehension
203 questions (Experiment 1). We show that, when context generates strong expectations
204 regarding the speaker's intentions (operationalized as an expected answer to a polar
205 question), understanding irony can be just as fast as - or even faster than - understanding a
206 literal reading of the same sentence. In Experiment 2, we additionally show that individual
207 differences among participants - with respect to their mindreading abilities - can account for
208 differences in irony comprehension. These results provide empirical support for the echoic
209 account and help explain the oft-reported variations in the literature with respect to the
210 processing effort of ironic readings of sentences relative to literal ones.

211 Experiment 1

212 With the idea of testing the echoic account of irony comprehension, we set up story
213 frames in such a way that a speaker's intention can be understood by the reader through two
214 channels: (i) by providing information about a speaker's informative intention with respect
215 to their audience (in the story) and; (ii) by providing information to the reader about the
216 speaker's attitude towards his or her own upcoming utterance. When this information is not
217 available, irony comprehension should not be facilitated. Let us consider these two pieces of
218 information in turn.

219 The first variable concerns the expectations that a reader is induced to have with
220 respect to the eventual speaker through the story situation. This can be illustrated again
221 through our opening example. In the 'strong expectation' context of the Guitar story (see
222 Figure 1), the reader is encouraged to expect the store owner to not agree to Chris's request.

223 This occurs through various pieces of information in the context, such as i) indicating that
224 the guitar is the most valuable in his shop, ii) explaining that the person who would handle
225 the guitar is a five-year-old, and iii) that the store owner dislikes children. Note that in the
226 Neutral condition, there are no such statements that serve as cues to the eventual speaker's
227 state of mind.

228 The second variable concerns explicit information about the speaker's attitude, which
229 indicates that the speaker's upcoming utterance is dissociated (insincere in some way) or
230 sincere. In the conditions that encourage dissociated attitudes towards the speaker's
231 upcoming utterance, readers will encounter statements such as *the owner has a reputation*
232 *for being a jokester, he therefore replies:* just before reading the speaker's actual utterance.
233 In the sincere conditions, which encourage readers to take the upcoming speaker's utterance
234 at face value, readers receive statements such as *the owner has a reputation for being frank,*
235 *he therefore replies:* as a lead up to the utterance.

236 As can be seen, the current design ultimately depends on a critical polar question.
237 Polar questions were chosen because they typically allow for two possible answers: 'yes' or
238 'no.' As far as irony inducing readings go, the polar question is useful because it arrives at a
239 moment in which readers can determine a) that the eventual speaker is likely to not comply
240 with the request (this is the strong expectation context) and that b) the eventual speaker
241 will reply with a dissociated attitude. As far as literal inducing readings go, there is little
242 intention-revealing information provided (this is the neutral context) and the eventual
243 speaker is described as speaking sincerely. With this design, the speed of comprehending the
244 target utterance can conceivably be fastest under conditions that optimize irony
245 understanding. For completeness, these two features are manipulated as part of a 2 x 2
246 design.

247 The current manipulation allows us to do two things. First, we can investigate the
248 effect of mindreading on irony comprehension on a trial-by-trial basis. Second, it will put us
249 in a position to directly determine whether the effect of mindreading can account for

250 differences in processing effort of ironic sentences relative to literal sentences.

251 **Participants and power analysis**

252 We wanted to determine the minimum number of participants that would allow us to
253 detect a true effect (more conservative in size than that found in the pilot, see supplementary
254 materials) with at least 80% power. To do this, we used the model parameters from the
255 analysis of the pilot study (i.e., the linear, mixed-effects model of the log-transformed
256 reaction times). These models had the following maximally-converging random effects
257 structure: The sum-contrast coded model that tested the interaction between both factors
258 included random intercepts by items and by participants. It also included random slopes for
259 both factors and their interaction by items. The sliding-contrast coded model included
260 random intercepts and slopes by items and random slopes by participants. This information
261 can be found in detail in the corresponding R script found on the project's OSF repository.

262 Crucially, we changed the estimated model coefficients for considerably more
263 conservative ones: We settled on an effect size with a Cohen's d value of 0.2 for all effects.
264 This is a more conservative estimate for every effect found in the pilot study and is commonly
265 used as a benchmark number for a 'small' effect size in psychological research (Cohen, 1992).
266 For the interaction effect, we settled on an effect size of half the size of the effect found in
267 the pilot study. Table 3 below summarizes the size of the relevant effects found in the pilot
268 study, the corresponding effect size used for computing power, the estimated statistical power
269 for finding such an effect with 220 participants, and the actual effect found in Experiment 1.
270 To estimate statistical power, we used a simulations-approach via the R package SimR
271 (Green & MacLeod, 2016). We simulated the results of 1000 experiments (for every relevant
272 effect) assuming the effect size shown in Table 1. We then counted the number of
273 experiments that found a significant effect, and used this number to estimate power.

Table 1

Effect sizes computed for the power analysis of Experiment 1. Effect sizes are given in Cohen's D.

Comparison	Effect size found in Pilot 1	Assumed effect size for simulations	Statistical power with 220 participants	Effect size found in Experiment 1
ATTITUDE*BIAS	0.397	0.2	86.4%	0.269
Interaction				
Sinc./neg. v. Insinc./neg.	0.389	0.2	89.1%	0.42
Insinc./neg. v. Insinc./neutral	0.6	0.2	81.0%	0.228
Sinc./neutral v. Insinc./neutral	0.4	0.2	87.0%	0.2

274 Participants recruited for the Experiment were right-handed, native speakers of
 275 American English between the ages of 18-35. In anticipation that some participants would
 276 not meet the exclusion criterion (correctly answering at least 5 out of 7 filler comprehension
 277 questions), we recruited a total of 319 participants via the online platform Prolific. Of these
 278 319, 57 (i.e., 17%) did not meet the inclusion criterion and were removed from the analysis,
 279 leaving the final number at 262. Participants gave their informed consent and received
 280 monetary compensation for their participation.

281 Design

282 Experiments 1 and 2 were programmed using the Ibex experimental software
283 (Drummond, 2013) coupled with the PennController (Zehr & Schwarz, 2018) and run via the
284 internet. Experiment 1 had a 2X2 design with the factors SPEAKER INTENTION (‘neutral’
285 vs. ‘strong expectation’) and SPEAKER ATTITUDE (‘sincere’ vs. ‘insincere’). All
286 manipulations refer to the type of contextual information that participants read prior to the
287 target utterance, which was always identical in every condition. Again, Figure 1 shows an
288 example critical item.

289 There were a total of nine critical items. For every participant, a new list was
290 automatically created showing only one out of the 4 possible versions of each critical item
291 (using Ibex’s built-in Latin-square design function). Because we had an odd number of 9
292 items, each participant saw 2 instances of three of the conditions and 3 instances of one of
293 the conditions. The condition for which participants saw one additional instance was
294 counterbalanced across participants. Participants also saw ten filler items. We settled on this
295 number of items for two reasons. First, since the experiment was to be web-based, it was
296 important to keep the task as short as possible to maintain participants’ attention and
297 minimize noise, following Futrell (2012). Second, we wanted to avoid any potential trial
298 effects, which have consistently been found to interact with processing effort of ironic relative
299 to literal sentences (Olkonemi, Ranta, & Kaakinen, 2016; e.g., Spotorno & Noveck, 2014).
300 Despite this low number of items, our a-priori power analysis showed that Experiment 1 was
301 sufficiently powered to find a true interaction effect (smaller than the one we actually found)
302 between SPEAKER INTENTION and SPEAKER ATTITUDE (see supplementary materials
303 for details on the power analysis).

304 There were comprehension questions after each critical item and after 7 out of the 10
305 filler trials. The critical comprehension questions assessed whether participants understood
306 the sentence ironically or literally (see Figure 1). The filler questions determined if
307 participants were included in the analysis or not: they had to answer at least 5 of the 7 filler

308 questions correctly. Filler and critical trials were pseudo-randomized, so that there would be
 309 at least one filler trial between every critical trial.

Factor 1: SPEAKER INTENTION		
Strong expectation	Neutral	
<i>Chris wants to buy his five-year-old daughter her first guitar. They go to a professional music shop together and she heads for the oldest and most valuable guitar in the store, which was behind a protective glass case. As she comes closer, one can see that the guitar is twice her size. The owner of the store, who really hates children, sees this and anxiously walks towards them. Chris sees him and says: "Sorry, could my daughter play this guitar?"</i>	<i>Chris wants to buy his 15-year-old daughter a new guitar, so they go to a music shop together. She is overwhelmed by all the different types of guitars they have, so she doesn't know which one to pick. They browse around for a while, and finally she finds one that she really likes, even though Chris doesn't understand why. He starts looking for the owner to ask him about it. Chris sees him and says: "Sorry, could my daughter play this guitar?"</i>	
Factor 2: SPEAKER ATTITUDE		
Insincere Attitude	Sincere Attitude	
<i>The owner has a reputation for being a jokester. He therefore replies:</i>	<i>The owner has a reputation for being frank. He therefore replies:</i>	
Target Sentence		
<i>"Yes, this guitar is here for everyone to play with"</i>		
Wrap-up Sentence		
<i>There were many other costumers in the store that day.</i>		
Comprehension question and possible answers		
<u>The owner will:</u>		
<i>(A) not let her play the guitar</i>	<i>(B) let her play the guitar</i>	<i>(C) Buy a guitar</i>

Figure 1. Example of a target utterance in Experiment 1 in the four conditions resulting from crossing the factors SPEAKER INTENTION and SPEAKER ATTITUDE. Note that Experiment 2 only had two conditions: 'strong expectation-insincere attitude' and 'neutral-sincere attitude'

310 **Materials**

311 Each critical item consisted of 8 sentences. The first five sentences set up expectations
 312 regarding the answer to the upcoming polar question (again, see Figure 1): Participants
 313 should strongly expect a 'no' answer, or not expect any particular answer whatsoever. These
 314 expectations were normed in a separate rating experiment, which is reported in the
 315 supplementary materials (<https://osf.io/vgkst/>). After these five sentences, participants read
 316 three additional ones: (1) a sentence that conveyed the attitude of the speaker and how it
 317 relates to the upcoming target sentence ('sincere attitude' or 'insincere attitude' conditions),
 318 (2) a target sentence that was always a 'yes' response and was identical across conditions,

319 and (3) a final wrap-up statement identical across conditions.

320 After each critical trial, participants chose one of three possible answers from a
321 multiple-choice question regarding the outcome of the situation. Their choice indicated
322 whether they constructed an ironic interpretation, a literal interpretation, or whether they
323 misunderstood the story altogether (i.e., a ‘distractor’ answer) (answers A, B, and C in
324 Figure 1 respectively). Position of the answers was randomized across trials.

325 **Procedure**

326 At the beginning of the experiment, participants were told that they were going to
327 read normal, every-day conversations and that they should imagine how these conversations
328 would play out in real life. They were not told that any of the exchanges were going to be
329 ironic. After completing two practice trials, the experiment began. Participants read all
330 stories on a sentence-by-sentence basis and hit the space-bar to reveal the next sentence.
331 When doing so, the previous sentence was replaced by dashed lines. For the comprehension
332 questions, participants could either use their keyboard (by pressing the numbers 1-3) or their
333 mouse to select one of the three possible answers. Participants took 9 minutes on average to
334 complete the Experiment.

335 **Predictions**

336 The predictions for both Experiment 1 and Experiment 2 were pre-registered. The
337 pre-registrations - along with all materials from both experiments, data and analysis scripts -
338 can be found on the project’s OSF page: <https://osf.io/329cs>

339 **Comprehension questions.** We reasoned that participants should be able to use
340 explicit information about a speaker’s attitude to understand whether an utterance is literal
341 or ironic. We therefore predicted that there should be a main effect of SPEAKER
342 ATTITUDE on comprehension, with items in the ‘insincere’ condition being taken as ironic
343 and those in the ‘sincere’ conditions as literal.

344 **Reading times.** If the type of context we created mediates processing effort of irony,
345 we should see that participants take less time reading ironic sentences (i.e. what we predict
346 to be sentences in the ‘insincere’ conditions) when there is a strong expectation compared to
347 when there is no expectation in particular. This should translate to a significant difference in
348 reading times between the ‘strong expectation-insincere attitude’ and the ‘neutral
349 context-insincere attitude’ conditions. Further, we predicted the opposite pattern for literal
350 sentences (the ‘sincere’ conditions): When participants expect the speaker to be sincere, they
351 should struggle processing a ‘yes’ response when they strongly expected a ‘no,’ whereas they
352 should have no difficulty reading the ‘yes’ response when they are not expecting any
353 particular answer (or arguably a ‘yes’ response by default). These differences should result in
354 a significant interaction between the two factors (SPEAKER INTENTION and SPEAKER
355 ATTITUDE).

356 **Analysis and results**

357 As a reminder, data from participants who did not answer at least 5 out of 7 of the
358 filler comprehension questions correctly were discarded, resulting in the exclusion of 57
359 participants (17%). Trials in which participants selected the distractor response (answer ‘C’
360 in Figure 1), were also discarded. This led to removing 3.4% of critical trials.

361 All remaining data was analyzed using the Lme4 package (Bates, Mächler, Bolker, &
362 Walker, 2015) in R (R Core Team, 2020). The data and analysis script for Experiment 1 are
363 available on the project’s OSF repository: <https://osf.io/vgkst/>. Models were fitted
364 following the recommendations of Barr, Levy, Scheepers, and Tily (2013). They included
365 random intercepts and slopes by items and participants for SPEAKER INTENTION,
366 SPEAKER ATTITUDE and their interaction, but excluded the random correlation between
367 intercept and slopes by participants.

368 **Comprehension questions.** Panel A of Figure 2 shows the resulting average
369 responses by condition. Target sentences in the insincere conditions were understood mostly

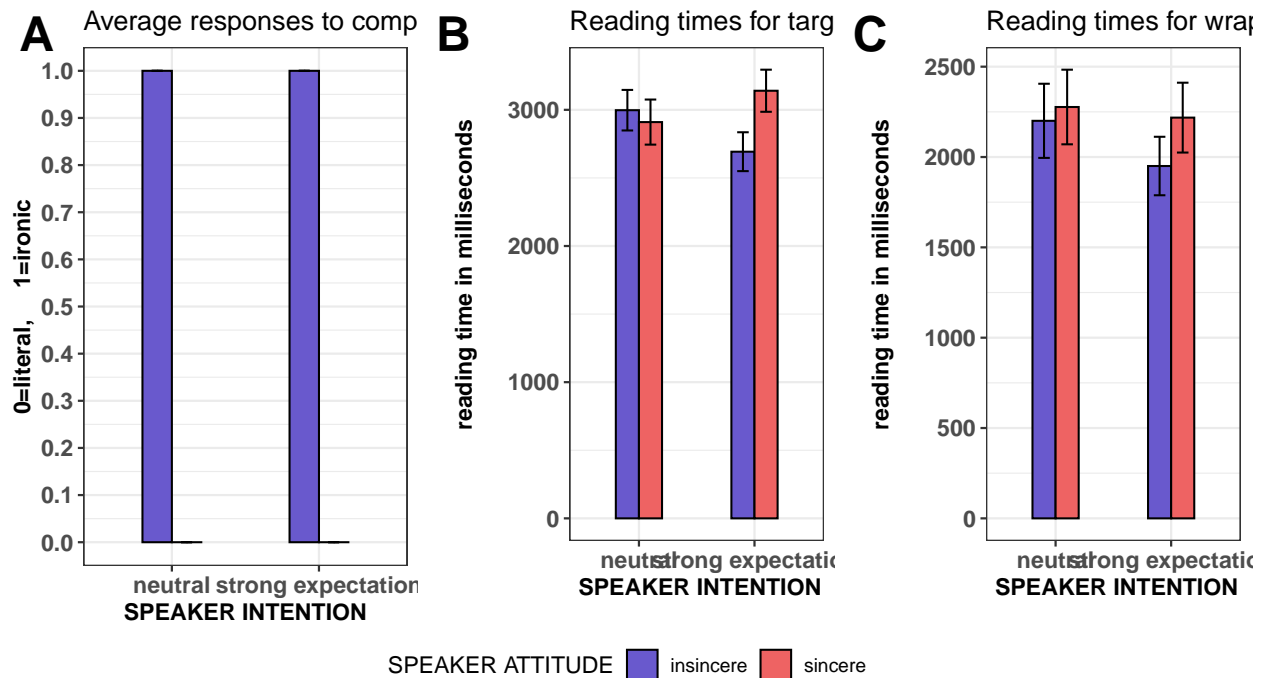


Figure 2. Average responses to comprehension questions (panel A) and raw-reading times of target (panel B) and wrap-up (panel C) sentences for trials with correct responses for Experiment 1. Error bars show confidence intervals.

370 as ironic (around 70% of the times), particularly in the strong expectation condition (82%).
 371 Sentences in the sincere conditions were perceived as literal (around 82% of the times),
 372 particularly in the neutral condition (around 93% of the times).

373 We fitted a mixed-effects logistic regression model to the data (sum-contrast coded).
 374 The reference levels for each factor were the neutral condition (factor: SPEAKER
 375 ATTITUDE) and the sincere condition (factor: SPEAKER INTENTION). The results
 376 confirmed our prediction and showed a main effect of SPEAKER ATTITUDE ($p < 0.001$,
 377 $z = 9.92$). There was an additional effect of SPEAKER INTENTION ($p = 0.001$, $z = -3.37$) and
 378 no significant interaction ($z = -0.21$, $p = 0.837$). The results are shown in Table 2. Overall,
 379 both ‘insincere’ conditions were understood above chance as ironic and both ‘sincere’
 380 conditions as literal.

Table 2

Summary of model output for accuracy in comprehension questions, Experiment 1

term	$\hat{\beta}$	95% CI	z	p
SPEAKER INTENTION	2.05	[0.86, 3.25]	3.37	.001
SPEAKER ATTITUDE	4.25	[3.41, 5.10]	9.92	< .001
ATTITUDE x BIAS interaction	-0.08	[-0.83, 0.68]	-0.21	.837

Note. model used a sum-contrast coding scheme

Table 3

Summary of model output for reading times of target sentence, Experiment 1

term	$\hat{\beta}$	95% CI	t	df	p
SPEAKER INTENTION	0.01	[-0.02, 0.05]	0.79	8.20	.453
SPEAKER ATTITUDE	-0.02	[-0.08, 0.03]	-0.85	7.96	.421
ATTITUDE x BIAS interaction	-0.08	[-0.12, -0.03]	-3.59	8.88	.006

Note. model used a sum-contrast coding scheme

Table 4

Comparison between reading times in individual conditions, Experiment 1

term	$\hat{\beta}$	95% CI	t	df	p
Sincere/strong e. vs. Insincere/strong e.	-0.20	[-0.26, -0.14]	-6.39	64.01	< .001
Insincere/strong e. vs. Insincere/neutral	0.11	[0.05, 0.18]	3.63	70.18	.001
Sincere/neutral vs. Insincere/neutral	-0.09	[-0.15, -0.02]	-2.75	67.09	.008

Note. model used a sliding-contrast coding scheme

Table 5

*Summary of model output for reading times of the wrap-up sentence,
Experiment 1*

term	$\hat{\beta}$	95% CI	t	df	p
SPEAKER INTENTION	-0.02	[-0.06, 0.02]	-0.93	6.82	.383
SPEAKER ATTITUDE	-0.04	[-0.07, 0.00]	-2.03	9.03	.073
ATTITUDE x BIAS interaction	-0.04	[-0.07, 0.00]	-1.83	7.35	.108

Note. model used a sum-contrast coding scheme

381 **Reading times.** To analyze reading times, we first excluded all incorrect responses.
 382 We then fitted a linear mixed-effects regression model to the log-transformed reading data of
 383 the target sentence. We settled on a log-transformation following the results of a box-cox
 384 test (Box & Cox, 1964). This test was performed because the residuals of a model using
 385 raw-reading times were not normally distributed.

386 The final model had an anova-style sum-contrast coding scheme, which allows us to
 387 test for main effects, and more importantly, it allows us the test the pre-registered prediction
 388 of their interaction directly. This model showed no main effects of SPEAKER ATTITUDE
 389 or of SPEAKER INTENTION. It did, however, show a significant interaction between both
 390 terms ($p=0.006$, $t=3.59$), in accord with our predictions. This model can be seen in Table 3.

391 With the thought of comparing reading times of ironic and literal interpretations of the
 392 same sentence, we followed up on these results by re-fitting the model using a sliding
 393 contrast coding scheme (as per our pre-registration). This form of contrast coding compares
 394 neighboring factor levels, which allows us to directly compare each relevant condition to each
 395 other. Specifically, we wanted to compare the two ‘insincere’ conditions (‘strong expectation’
 396 and ‘neutral’) to one another, the two ‘strong expectation’ conditions (‘sincere’ and
 397 ‘insincere’) to one another, and the two ‘neutral’ conditions (‘sincere’ and ‘insincere’) to one
 398 another. This new model showed a significant difference between ‘sincere-strong expectation’

399 and ‘insincere-strong expectation’ conditions ($p < 0.001$, $t = 6.39$), a significant difference
400 between ‘insincere-strong expectation’ and ‘insincere-neutral’ ($p = 0.001$, $t = 3.63$), and a
401 significant difference between ‘sincere-neutral’ and ‘insincere-neutral’ conditions ($p = 0.008$,
402 $t = 2.75$). This model can be seen in Table 4. There were no spill-over effects found in the
403 wrap-up sentence (see panel C of Figure 2 and Table 5 for the summarized results).

404 Discussion

405 Experiment 1 manipulated two sorts of information put at a participant’s disposal prior
406 to hearing a potentially ironic remark. These corresponded with two aspects of mindreading,
407 namely (i) the degree to which information in the context allows a reader to anticipate a
408 speaker’s intention and (ii) explicit information about the speaker’s attitude towards their
409 own utterance. These two aspects are central to the echoic account of irony comprehension
410 (Wilson & Sperber, 2012). Overall, our results showed that both (i) and (ii) affected
411 comprehension of our target utterances. This pattern played out differently in reading times
412 than it did in interpretation: For the ultimate interpretation of the sentence (quantified as
413 responses to the comprehension question), speaker-specific cues about a speaker’s attitude
414 towards their upcoming utterance was the most important factor, with both sincere
415 conditions being mostly understood as literal and both insincere conditions as ironic.
416 However, the degree to which it is possible to anticipate a speaker’s intention also influenced
417 participants’ irony comprehension: The more a participant expected a ‘no’ answer, the more
418 they understood a ‘yes’ answer as ironic. This resulted in two main effects and no interaction.

419 For reading times, on the other hand, the interaction between both types of cues was
420 crucial: when a sentence was understood as ironic (‘insincere’ conditions), it was read faster
421 if participants had strong intuitions regarding the informative intention of the speaker
422 (‘insincere-strong expectation’ condition) compared to when they did not (‘insincere-neutral’
423 condition). This finding supports the idea that differences in mindreading engagement
424 (operationalized here as the degree to which a context allows a participant to anticipate the

425 speaker's upcoming intention as well as attitude towards a proposition) predict ease of
426 processing ironic sentences.

427 A closer look at this interaction effect has an important bearing on the "ironic
428 vs. literal" debate. First, consider the comparison of the 'insincere-strong expectation'
429 condition to the 'sincere-strong expectation' condition. Among these two in the strong
430 expectation condition, the one encouraging an ironic reading of a sentence is actually faster.
431 Second, consider the 'sincere-neutral' condition as it is compared to the 'insincere-neutral'
432 condition. Here, the ironically understood sentences are read slower than their literal
433 counterparts. These findings therefore suggest two things. First, there is no primacy of the
434 literal meaning regarding the processing speed of an entire sentence: We failed to find a main
435 effect of ATTITUDE, which suggests that literal sentences were not faster to process than
436 ironic ones across the board. Second, the underlying factor that mediates processing effort
437 might not be whether the sentence is ironic or literal, but the degree to which context gives
438 readers access to the speaker's intentions and beliefs. These results provide evidence for a
439 likely rapid engagement of mindreading abilities when understanding irony. It also makes for
440 a very rare finding of irony understanding actually being faster than its explicitly literal
441 control. In Experiment 2 we seek to find further support for our claim by investigating
442 individual differences between participants in comprehending ironic and literal sentences.

443

Experiment 2

444 Experiment 1 showed that mindreading considerations mediate irony comprehension
445 effort. We view this as being a consequence of how participants use their mindreading skills:
446 Having access to a speaker's intention can predict the comprehender's ease of irony
447 processing and comprehension accuracy. However, it could be the case that participants in
448 Experiment 1 were not engaging their mindreading skills, but instead learning to associate
449 specific lexical cues in the context with a potential interpretation and used this association
450 as a comprehension strategy. In other words, it could be that participants relied on

451 contextual cues without considering the speaker's intentions. To support our interpretation
452 of Experiment 1, we need to seek out evidence suggesting that mindreading is at play.

453 We decided to go about this by taking an individual differences approach. Apperly
454 (2012) argued that there are individual differences with regards to the degree to which people
455 can routinely and appropriately put their mindreading skills to use. This has been shown to
456 have repercussions for pragmatic language comprehension, in as much as people with more
457 developed mindreading skills tend to show a better understanding of various pragmatic
458 phenomena such as irony (Spotorno & Noveck, 2014), scalar implicatures (Fairchild &
459 Papafragou, 2021) and humor (Bischetti, Ceccato, Lecce, Cavallini, & Bambini, 2019). If the
460 differences between conditions in Experiment 1 were linked to differences in mindreading
461 engagement, we should be able to find an association between comprehension of the critical
462 items of Experiment 1 with individual differences in mindreading abilities. This is the goal of
463 Experiment 2.

464 **Participants and power analysis**

465 To calculate power for Experiment 2, we ran a power analysis via simulations, similar
466 to the procedure of Experiment 1. The main difference here is that the effect of interest for
467 Experiment 2 was the interaction in the logistic regression model. The model used for the
468 simulations included a maximally-converging random effects structure of random slopes by
469 items and random intercepts by participants. Since it is not possible to calculate Cohen's D
470 for a logistic regression model, we used a conservative estimate of half of the raw-effect size
471 found in the pilot (i.e., the beta coefficient of the interaction term, see Table 2 of the
472 supplementary materials). After simulating 1000 Experiments using the pilot's parameters
473 and this new - conservative - beta coefficient, we concluded that an Experiment with 220
474 participants would have over 80% power to detect a true effect of that magnitude. The final
475 effect found in Experiment 2 was larger than this conservative estimate, showing that
476 Experiment 2 was sufficiently powered. The power analysis and pilot data are available on

477 the project's OSF repository.

478 We thus recruited 239 participants (who did not participate in Experiment 1),
479 assuming that some might not meet the exclusion criterion: As in Experiment 1, we intend
480 to exclude participants who do not correctly answer at least 5 out of the 7 comprehension
481 questions in the filler items. For Experiment 2, the exclusion criterion led to the exclusion of
482 16 participants, leaving the final number at 223.

483 **Materials, design and procedure**

484 The materials, design and procedure were similar to that of Experiment 1. There were
485 three differences: First, we kept only the 'insincere - strong expectation' and the 'sincere -
486 neutral' conditions, since these two conditions were the ones that were most typically
487 understood as ironic and literal, respectively. Second, we decided to show participants 8 of
488 the critical items of Experiment 1 in these two conditions (i.e., 4 items in each condition).
489 This was done to balance the number of items in each condition seen by participants relative
490 to Experiment 1. Third (and most importantly), Experiment 2 included a mindreading task,
491 administered to participants after completing the experiment. This task was an abridged
492 version of the 'Reading the Mind in the Eyes' (RME) task (Baron-Cohen, Wheelwright, Hill,
493 Raste, & Plumb, 2001), meant to measure each participant's ability to deploy their
494 mindreading skills. This abridged version consisted of the first 24 trials of the task. We chose
495 to use an abridged version in order to keep the experiment as a whole as short as possible.
496 Everything else was identical to the original task by Baron-Cohen, Wheelwright, Hill, Raste,
497 and Plumb (2001). We opted for the RME instead of other advanced mindreading measures
498 such as the 'Strange Stories' task (F. G. Happé, 1994) because the former relies less than the
499 latter on pragmatic competence, i.e. on understanding language use in specific contexts
500 (Bosco, Tirassa, & Gabbatore, 2018). As Bosco, Tirassa, and Gabbatore (2018) argue, when
501 tasks explicitly rely on figurative language comprehension and pragmatic inferencing as
502 measures of higher mindreading abilities (such as the 'Strange Stories' does), it is difficult to

503 estimate the true degree to which mindreading correlates with the comprehension of
504 pragmatic phenomena (such as irony), given that both things are effectively measured with
505 the same task. We address the limitations of using the RME task in the General Discussion.

506 We computed a mindreading score for each participant based on their results on the
507 RME task. This score was used as a continuous predictor for analyzing the responses to the
508 comprehension questions and the reading times of the target sentence. Together with this
509 continuous predictor (which we refer to as MINDREADING), we coded the ‘insincere -
510 strong expectation’ and the ‘sincere - neutral’ conditions as two levels (‘ironic’ and ‘literal,’
511 respectively) of the same factor (SENTENCE TYPE) and included them as predictors of
512 comprehension accuracy and response times. We also included the interaction between
513 MINDREADING and SENTENCE TYPE as a predictor.

514 **Predictions**

515 The landmark study by F. G. E. Happé (1993) showed that irony comprehension
516 correlated with success in a second-order false-belief task, which led her to interpret the
517 results as supporting the echoic account. Wilson and Sperber (2012) (pg. 134) echo this
518 interpretation by stating that Happé’s results “confirm the relevance-theoretic account of
519 figurative utterances.” We interpret this as an indicator that the echoic account predicts that
520 mindreading scores should correlate necessarily with irony comprehension scores. However,
521 the theory seems to remain vague as to whether mindreading scores should also correlate
522 with irony processing speed. For this reason, our pre-registered predictions refer to sentence
523 comprehension only, as indicated by responses to the comprehension questions after the
524 critical items, and we analyze the reading time data as an exploratory measure only.

525 We hypothesized that if the context cues used in Experiment 1 (information about the
526 speaker’s attitude towards their upcoming proposition and a contextual bias towards
527 expecting a ‘no’ answer to the polar question) reflect the way in which comprehenders engage
528 in mindreading abilities, there should be a link between an individual comprehender’s level

529 of mindreading skill and their responses in the different conditions (specifically, the ‘insincere
 530 - strong expectation’ and the ‘sincere - neutral’ conditions of Experiment 1, which are called
 531 ‘ironic’ and ‘literal’ in Experiment 2). This should result in a significant interaction between
 532 MINDREADING and SENTENCE TYPE for responses to the comprehension questions.

533 Concretely, we predicted that participants with higher mindreading scores should be
 534 better at understanding irony in the ‘ironic’ condition compared to participants with lower
 535 mindreading scores. No such effect of MINDREADING should be visible in the ‘literal’
 536 condition. This prediction reflects that (i) we believe the pattern of results of Experiment 1
 537 to be related to mindreading engagement, and (ii) enhanced mindreading should be
 538 particularly advantageous for understanding ironic utterances and not their literal
 539 counterparts. These predictions directly motivate how we analyzed our data, which we
 540 describe in the following section.

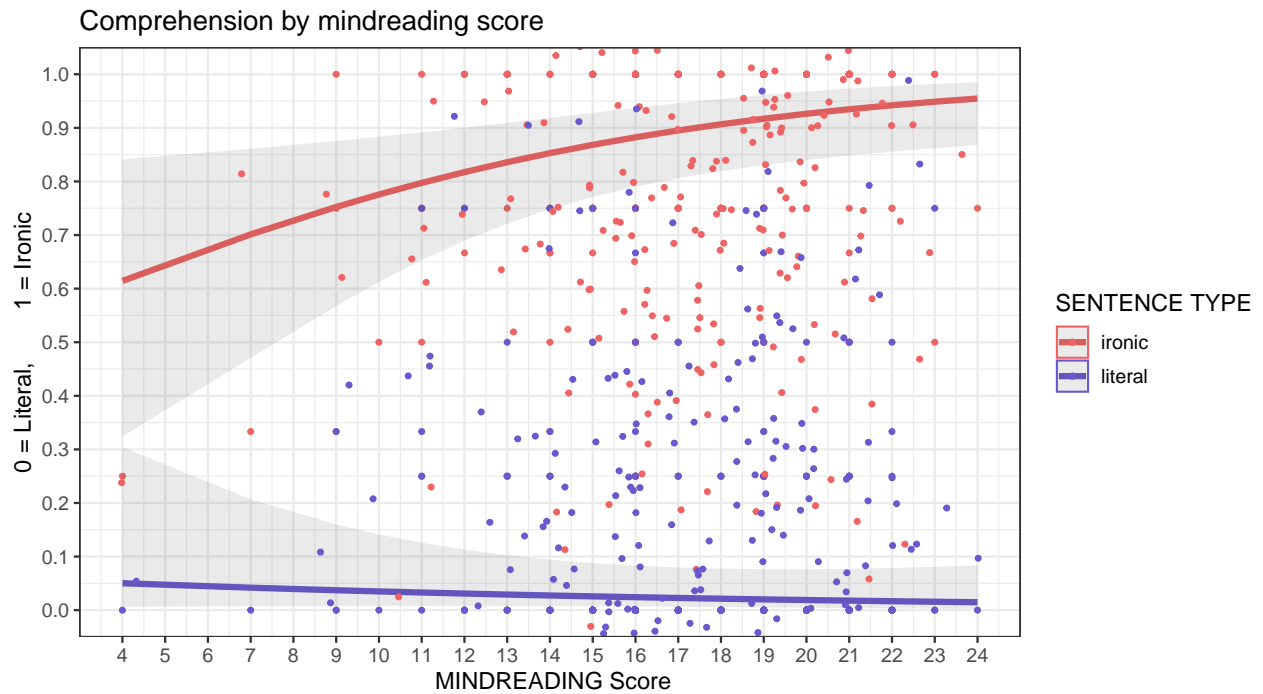


Figure 3. Responses by mindreading scores in Experiment 2. Individual dots show participant averages. Plotted lines and gray ribbons show the predicted values of the logistic regression model and confidence intervals, respectively.

Table 6

Model results for comprehension questions (ironic), Experiment 2

term	$\hat{\beta}$	95% CI	z	p
MINDREADING	0.13	[0.03, 0.23]	2.59	.010
SENTENCE TYPE	-2.64	[-5.59, 0.32]	-1.75	.081
MINDREADING x SENTENCE TYPE interaction	-0.19	[-0.36, -0.03]	-2.28	.023

Note. model used a treatment-contrast coding scheme, ironic condition is coded as the baseline

541 **Analysis and results**

542 **RME task.** Because we used an abridged version of the RME task, we assessed our
 543 version’s internal consistency to evaluate its similarity to the original, un-abridged version.
 544 We found that our task had a Cronbach’s alpha value of 0.61, similar to that found in
 545 previously reported uses of the unabridged version (e.g., Harkness, Sabbagh, Jacobson,
 546 Chowdrey, & Chen, 2005; Vellante et al., 2013; Voracek & Dressler, 2006). We also found a
 547 McDonald’s Omega value of 0.64. Additionally, an exploratory factor analysis suggests that
 548 a unidimensional model was an adequate fit to the data, with an RMSEA index of 0.041 and
 549 a BIC of -987. A model with three factors was a better fit to the data (BIC: -898), a
 550 phenomenon which has also been observed for the unabridged version (Olderbak et al., 2015).
 551 These analyses can be found in the supplementary materials.

552 **Comprehension questions.** After excluding trials in which participants selected
 553 the distractor response (4% of all trials), we fitted a mixed effects logistic regression model
 554 to analyze comprehension data, as we did in Experiment 1. In Experiment 2, the model
 555 included the factor SENTENCE TYPE (levels: ‘ironic’ and ‘literal’), the continuous
 556 predictor MINDREADING (which was first scaled), and their interaction. Since the goal of
 557 our analysis was to test the interaction and see whether MINDREADING affected the two
 558 levels of SENTENCE TYPE differently (as per our pre-registered predictions), we fitted the

Table 7

Summary of model output for comprehension questions (literal), Experiment 2

term	$\hat{\beta}$	95% CI	z	p
MINDREADING	-0.05	[-0.14, 0.04]	-1.05	.293
SENTENCE TYPE	1.68	[-0.50, 3.86]	1.51	.131
MINDREADING x SENTENCE TYPE interaction	0.19	[0.07, 0.31]	2.99	.003

Note. model used a treatment-contrast coding scheme, literal condition is coded as the baseline

559 same model twice using treatment contrast coding: One in which the ‘ironic’ condition was
 560 coded as the baseline, and another in which the ‘literal’ condition was coded as the baseline.
 561 When using a treatment contrast-coding scheme, the coefficients of each predictor represent
 562 an effect relative to the baseline condition only. In other words, with treatment contrast we
 563 only test the simple effect of MINDREADING on the baseline condition (‘ironic’ and ‘literal,’
 564 in each of the two models), instead of the main effect of MINDREADING on responses
 565 across conditions. Re-fitting the model thus addresses the prediction that MINDREADING
 566 should impact irony comprehension but not the comprehension of literal sentences. Both
 567 iterations of the model included random intercepts and slopes for MINDREADING,
 568 SENTENCE TYPE and their interaction by items, and a random intercept and slope term
 569 for SENTENCE TYPE by participants.

570 The results of the model are summarized in Table 6, and the results pattern is
 571 illustrated in Figure 3. As predicted, there was a significant interaction between
 572 MINDREADING and SENTENCE TYPE ($z=2.28$, $p=0.023$). This suggests that
 573 MINDREADING had a different effect on each of the levels of SENTENCE TYPE: As
 574 predicted, higher mindreading scores resulted in significantly more correct interpretations in
 575 the ‘ironic’ condition ($z=2.59$, $p=0.01$), and we failed to find an effect of MINDREADING
 576 on the ‘literal’ condition ($z=1.05$, $p=0.293$) (see Table 7).

577 **Reading times.** As in Experiment 1, we first excluded incorrect responses from the
578 analysis (i.e., ironic answers in the ‘literal’ condition and ‘literal’ answers in the ironic
579 condition). This resulted in the removal of 14% of the data. We fitted a mixed-effects linear
580 regression model to the log-transformed reading times of remaining data. We included
581 MINDREADING, SENTENCE TYPE, and their interaction as predictors. The final
582 random-effects structure included random intercepts by participants and items, as well as a
583 random slope term for SENTENCE TYPE by items. We failed to find any significant effects
584 of our predictors on the log-transformed reading times.

585 **Discussion**

586 In Experiment 2, we anticipated that individual participants’ scores on the ‘Reading
587 the Mind in the Eye’ task would be predictive of their accuracy in understanding irony - but
588 not in understanding literal sentences. The results confirmed our predictions: Participants
589 with higher mindreading scores were better at understanding irony than those with lower
590 mindreading scores, but not at understanding literal sentences. This result suggests that the
591 contextual manipulations of Experiment 1 - being aware of the speaker’s intention and
592 knowing the speaker’s attitude - were in fact tapping into the way in which participants
593 engaged their mindreading abilities during reading comprehension. This is true of at least
594 the ‘insincere-strong expectation’ and ‘sincere-neutral’ conditions of Experiment 1, which
595 were the most prototypically ironic and literal, respectively.

596 The failure to find effects of MINDREADING on reading times could have various
597 explanations. First, irony comprehension is quite low for participants in the bottom-half of
598 the distribution of RME scores. This means that there are very few instances of successful
599 irony comprehension for which we could measure RTs for these participants. It could be the
600 case that more observations are necessary to detect effects in this regard. However, precisely
601 because comprehension accuracy of irony is low for low-scores on the RME task, it is not
602 clear whether it would even be meaningful to interpret the processing effort of the instances

603 that these participants do accurately recognize as ironic. These might be either chance
604 occurrences or guided by altogether different comprehension mechanisms.

605 **General Discussion**

606 The goal of the current study was twofold. First, we aimed to investigate whether
607 providing contextual cues that point to a speaker's intention and their attitude towards an
608 upcoming proposition could explain reading times differences with respect to ironic relative
609 to literal readings (Experiment 1). Second, we wanted to examine whether any such irony
610 comprehension differences brought on by context were related to individual differences in
611 participants' engagement in mindreading abilities (Experiment 2).

612 Our results broadly support our predictions. First, having access to the speaker's
613 beliefs and intentions plus information about a speaker's attitude towards an upcoming
614 proposition provide the means for a rapid interpretation of an ironic response: When
615 participants were told that the speaker might not be committed to the truth of an utterance
616 (because he is 'known to be a jokester,' for example) they overwhelmingly understood the
617 target sentence as ironic compared to when they believed the speaker to be committed to the
618 truth of their utterance (Experiment 1). This finding is in line with previous research
619 showing that speaker-specific information affects the overall rate of interpretation of ironic
620 sentences (Katz & Pexman, 1997; Pexman & Olineck, 2002).

621 Second, our results show that encouraging participants to engage with a speaker's
622 intentions facilitates processing effort of utterances that are understood ironically. When
623 sentences were understood ironically (based on the responses to comprehension questions),
624 participants read them faster if they were embedded in a context that made manifest a
625 specific intention (prior to a polar question) compared to when the discourse context did not
626 aim to generate any specific expectations. This finding complements the literature on the
627 interaction between mindreading skills and language processing (e.g., Ferguson & Breheny,
628 2011; Rubio-Fernández, Mollica, Oraa Ali, & Gibson, 2019) by showing a further

629 phenomenon for which mindreading, when engaged via a linguistic context, has a rapid effect.
630 The finding also supports and extends the work of Spotorno and Noveck (2014) by showing
631 that mindreading is critical for irony comprehension. Here, we showed how mindreading can
632 have an impact within individual trials; it need not rely on trial effects.

633 Third, we showed that our experimental manipulation correlated with individual
634 differences in mindreading abilities. Participants who scored lower on a mindreading test
635 were less accurate in understanding irony. This was not the case for the comprehension of
636 literal sentences, supporting the claim that mindreading skills are particularly relevant for
637 understanding ironically - but not literally - intended sentences.

638 Finally, and most importantly, our results contribute to the debate on processing ironic
639 relative to literal sentences. We show that knowledge about a speaker's attitude and
640 expectations about a speaker's intention interact during reading. The result of this
641 interaction is that one and the same sentence can be read faster or slower - and ironically or
642 literally - depending on the degree to which participants think they can anticipate the
643 intention and attitude of the speaker. We suggest that the key to understanding the
644 relationship between context and processing effort of ironic sentences is to focus on how a
645 context helps comprehenders anticipate the intentions of a speaker. With this in mind, we
646 can make sense of previous incongruous empirical findings: whenever a specific cue aids in
647 engaging a comprehender's mindreading abilities, it will ease processing, so that an ironic
648 interpretation can be reached just as fast as a literal one would.

649 **Interpreting the results of the RME task**

650 The RME has often been used as a measure of individual differences in mindreading
651 skills for neurotypical populations (e.g., Domes, Heinrichs, Michel, Berger, & Herpertz, 2007;
652 Kidd & Castano, 2013; Mar, Oatley, Hirsh, Dela Paz, & Peterson, 2006). However, the task's
653 validity and interpretation have been heavily criticized (e.g., Baker, Peterson, Pulos, &
654 Kirkland, 2014; Black, 2019; Oakley, Brewer, Bird, & Catmur, 2016). Oakley, Brewer, Bird,

655 and Catmur (2016) noted that performance in this task is likely driven by the ability to
656 identify emotional states rather than the ability to attribute mental states to others. The
657 authors argue that the RME task picks up on differences between ASD and neurotypical
658 populations because Autism often co-occurs with Alexithymia, a clinical condition
659 characterized by difficulties in identifying and describing one's emotional states and those of
660 others (see also Bird & Cook, 2013). This co-occurrence is so prevalent that difficulties in
661 the emotional domain have traditionally been considered a common trait of ASD and are
662 even used as diagnostic markers (Lord et al., 2000). This is important to consider given the
663 pattern of results of Experiment 2. Here, we see that the rate of irony comprehension as a
664 function of MINDREADING scores only seems to change for participants who scored in the
665 bottom-half on the RME task. This range of scores is well within the range typically found
666 for ASD individuals (mean=62.7%, according to a meta-analysis by Peñuelas-Calvo, Sareen,
667 Sevilla-Llewellyn-Jones, & Fern'andez-Berrocal, 2019). It is possible that at least some of our
668 participants displayed difficulties both in identifying emotions (as tracked by the RME task)
669 and in attributing mental states to others (since these two difficulties often co-occur for ASD
670 individuals). This would explain the differences in irony comprehension for participants with
671 low RME scores. However, we did not collect information on participants' ASD diagnosis, so
672 this remains speculative.

673 This interpretation is nonetheless consistent with some of the other criticism that the
674 RME task has received. For example, Black (2019) states that the RME task should be seen
675 as an instrument best suited for detecting strong mindreading differences such as those
676 between ASD and neurotypical populations. This could explain why, for the high performers
677 in the RME task of our Experiment 2, there was no obvious effect of RME scores on irony
678 comprehension: The RME task was likely not sensitive enough to detect any differences in
679 participants with normal to high mindreading abilities. In other words, the RME task might
680 be best seen as a sort of 'blunt' tool that can pick up on the substantial differences between
681 individuals with low- and high-mindreading abilities, but is not ideal for detecting the finer

682 differences between mindreaders at the upper end of the scale. The goal of Experiment 2 was
683 to determine whether the elicited differences in comprehension found in Experiment 1 could
684 be said to be related to the engagement of mindreading abilities. We interpret the broad
685 differences detected by the RME as sufficient evidence of this. However, future studies
686 interested in fine-grained differences among individuals with high mindreading skills should
687 rely on other measurement tasks better suited to that particular population.

688 **Implications for theories of irony comprehension**

689 The present work derived its predictions from the echoic account of irony (Wilson &
690 Sperber, 2012). We see our study as an extension of this account by way of providing
691 testable linking hypotheses for the theory. Specifically, we provide a principled explanation
692 of what type of context influences irony processing and why: A context that allows one to
693 have access to the speaker's likely informative intention and manifest access to the attitude
694 attached to their proposition will ease irony comprehension. The fact that the contexts we
695 created in Experiment 1 triggered an engagement of these two aspects of mindreading is
696 supported by the results of Experiment 2. Here, we found that individual differences in
697 mindreading abilities critically interact with the experimental conditions, at least in as much
698 as there is a difference between low and high-scorers on the RME task.

699 The results of the two experiments could also be interpreted as being compatible with
700 other theoretical views. Indeed, our findings can be considered compatible with contextualist
701 accounts such as the Direct Access View (Gibbs, 2002) and the Constraint-Satisfaction
702 Account (Pexman, 2008; Pexman, Ferretti, & Katz, 2000) in the sense that they underline
703 the role played by context during online processing of irony and show that irony can be
704 processed faster than literal language (under certain circumstances). However, a major
705 drawback of these accounts is the absence of a systematic weighing of contextual factors with
706 regards to how they affect processing. This makes it hard for the accounts to predict which
707 contextual aspects will facilitate irony processing and which will not.

708 Our findings are less compatible with context-independent accounts - such as the
709 Graded Salience Hypothesis (Giora, 2003; Giora et al., 2007) - since such accounts would
710 posit a primacy of salient readings at the sentence-level (which are literal in the majority of
711 cases), regardless of context. In Experiment 1, we showed how one and the same sentence
712 can be read faster or slower as a function of contextually raised expectations of a specific
713 nature. This resulted in irony sometimes being faster and sometimes slower than a literal
714 control, which is at odds with a stronger version of the context-independent view (for
715 example, as formulated by Schwoebel, Dews, Winner, & Srinivas, 2000) that would preclude
716 irony from being read faster than literal equivalents, regardless of a contextual bias. That
717 said, the goal of our study was not to test the predictions of the Graded Salience Hypothesis,
718 but to derive and test predictions from the echoic account in order to investigate the role
719 that mindreading plays with regards to irony comprehension. To test the predictions of the
720 Graded Salience Hypothesis, it would be necessary to carefully norm the target ironic
721 sentences for frequency, prototypicality, familiarity, and other factors that might affect
722 salience. We leave it to future work to study how a mindreading-facilitating context can be
723 used to test the predictions of context-independence views.

724 An alternative explanation of our results could relate to the granularity of our
725 measures. For example, it could be the case that participants in Experiment 1 first read the
726 initial part of the target sentence (specifically, the word ‘yes’) faster when intended literally
727 (i.e., the ‘sincere’ conditions) than when intended ironically (i.e., the insincere conditions)
728 and that a processing advantage for ironic readings only appeared later downstream as the
729 sentence was integrated with context. Though our data cannot rule out this scenario (since
730 we only measured reading times of the entire sentence), we think it is not likely given the
731 size of differences in reading times. For example, the effect size we found for the difference
732 between the Sincere/strong expectations (understood literally) vs. Insincere/strong
733 expectations (understood ironically) in Experiment 1 was of Cohen’s $d = 0.42$, or about 450
734 milliseconds advantage for the ironic condition. This is substantial, and it seems unlikely

735 that it would be this big if there had first been an effect in the opposite direction at the
736 beginning of the sentence.

737 Further, such a pattern would not be predicted by theoretical accounts that posit an
738 overall processing advantage for literal over ironic readings (as far as we can tell). For
739 example, the Graded Salience Hypothesis (probably the most prominent
740 ‘context-independence’ theory of irony) explicitly states that there should be no differences
741 between irony and literal readings of the same expressions at the earliest stages (i.e., at the
742 word-level processing stage), and that a processing speed advantage for salient-based (i.e.,
743 literal) readings should appear in a second stage at the earliest (Giora & Fein, 1999; Giora et
744 al., 2007, pg. 141; Giora, Fein, & Schwartz, 1998). Because of this, we see our explanation as
745 being more parsimonious: The degree to which participants can anticipate the speaker’s
746 intention is what drives processing speed differences between ironic and literal
747 interpretations of the same sentence.

748 **Conclusion**

749 The current study suggests that differences in mindreading engagement induced via a
750 discourse context result in systematic differences in irony comprehension. These results
751 provide linking hypotheses for the echoic account of irony comprehension as well as
752 contribute to the long-standing debate on the processing effort of verbal irony by showing
753 under which conditions ironic sentences can be read faster than literal equivalents.

References

754

755

756

Allott, N. (2017). Metarepresentation. In *The Routledge Handbook of Pragmatics* (pp. 295–309). Routledge.

757

758

759

Apperly, I. A. (2012). What is “theory of mind?” Concepts, cognitive processes and individual differences. *Quarterly Journal of Experimental Psychology*, *65*(5), 825–839.

760

761

762

Baker, C. A., Peterson, E., Pulos, S., & Kirkland, R. A. (2014). Eyes and IQ: A meta-analysis of the relationship between intelligence and “Reading the Mind in the Eyes.” *Intelligence*, *44*, 78–92.

763

764

765

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a “theory of mind”? *Cognition*, *21*(1), 37–46.
[https://doi.org/10.1016/0010-0277\(85\)90022-8](https://doi.org/10.1016/0010-0277(85)90022-8)

766

767

768

769

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The “Reading the Mind in the Eyes” Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *42*(2), 241–251.

770

771

772

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*(3), 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>

773

774

775

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

- 776 Bird, G., & Cook, R. (2013). Mixed emotions: The contribution of alexithymia to the
777 emotional symptoms of autism. *Translational Psychiatry*, 3(7), e285–e285.
- 778 Bischetti, L., Ceccato, I., Lecce, S., Cavallini, E., & Bambini, V. (2019). Pragmatics
779 and theory of mind in older adults' humor comprehension. *Current Psychology*,
780 1–17.
- 781 Black, J. E. (2019). An IRT analysis of the reading the mind in the eyes test. *Journal*
782 *of Personality Assessment*, 101(4), 425–433.
- 783 Bosco, F. M., Tirassa, M., & Gabbatore, I. (2018). Why Pragmatics and Theory of
784 Mind Do Not (Completely) Overlap. *Frontiers in Psychology*, 9.
- 785 Box, G. E. P., & Cox, D. R. (1964). An Analysis of Transformations. *Journal of the*
786 *Royal Statistical Society. Series B (Methodological)*, 26(2), 211–252.
- 787 Bryant, G. A. (2012). Is Verbal Irony Special?: Is Verbal Irony Special? *Language*
788 *and Linguistics Compass*, 6(11), 673–685. <https://doi.org/10.1002/lnc3.364>
- 789 Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155.
- 790 Degen, J., & Tanenhaus, M. K. (2019). Constraint-based pragmatic processing. *The*
791 *Oxford Handbook of Experimental Semantics and Pragmatics*, 21–38.
- 792 Domes, G., Heinrichs, M., Michel, A., Berger, C., & Herpertz, S. C. (2007). Oxytocin
793 improves “mind-reading” in humans. *Biological Psychiatry*, 61(6), 731–733.
- 794 Drummond, A. (2013). Ibex farm. *Online Server: Http://Spellout. Net/Ibexfarm.*
- 795 Fairchild, S., & Papafragou, A. (2021). The Role of Executive Function and Theory
796 of Mind in Pragmatic Computations. *Cognitive Science*, 45(2), e12938.

- 797 Ferguson, H. J., & Breheny, R. (2011). Eye movements reveal the time-course of
798 anticipating behaviour based on complex, conflicting desires. *Cognition*, *119*(2),
799 179–196. <https://doi.org/10.1016/j.cognition.2011.01.005>
- 800 Filik, R., Howman, H., Ralph-Nearman, C., & Giora, R. (2018). The role of
801 defaultness and personality factors in sarcasm interpretation: Evidence from
802 eye-tracking during reading. *Metaphor and Symbol*, *33*(3), 148–162.
803 <https://doi.org/10.1080/10926488.2018.1481258>
- 804 Filik, R., & Moxey, L. M. (2010). The on-line processing of written irony. *Cognition*,
805 *116*(3), 421–436.
- 806 Futrell, R. (2012). *Processing effects of the expectation of informativity*.
- 807 Gibbs, R. (1986). On the psycholinguistics of sarcasm. *Journal of Experimental*
808 *Psychology: General*, *115*(1), 3–15. <https://doi.org/10.1037/0096-3445.115.1.3>
- 809 Gibbs, R. (1994). *The Poetics of Mind: Figurative Thought, Language, and*
810 *Understanding*. Cambridge University Press.
- 811 Gibbs, R. (2002). A new look at literal meaning in understanding what is said and
812 implicated. *Journal of Pragmatics*, *34*(4), 457–486.
813 [https://doi.org/10.1016/S0378-2166\(01\)00046-7](https://doi.org/10.1016/S0378-2166(01)00046-7)
- 814 Giora, R. (2003). *On Our Mind: Salience, Context, and Figurative Language*. Oxford
815 University Press.
- 816 Giora, R., Drucker, A., Fein, O., & Mendelson, I. (2015). Default sarcastic
817 interpretations: On the priority of nonsalient interpretations. *Discourse Processes*,
818 *52*(3), 173–200.

- 819 Giora, R., & Fein, O. (1999). Irony: Context and salience. *Metaphor and Symbol*,
820 *14*(4), 241–257.
- 821 Giora, R., Fein, O., Laadan, D., Wolfson, J., Zeituny, M., Kidron, R., . . . Shaham, R.
822 (2007). Expecting irony: Context versus salience-based effects. *Metaphor and*
823 *Symbol*, *22*(2), 119–146.
- 824 Giora, R., Fein, O., & Schwartz, T. (1998). Irony: Grade salience and indirect
825 negation. *Metaphor and Symbol*, *13*(2), 83–101.
- 826 Green, P., & MacLeod, C. J. (2016). SIMR: An R package for power analysis of
827 generalized linear mixed models by simulation. *Methods in Ecology and Evolution*,
828 *7*(4), 493–498. <https://doi.org/10.1111/2041-210X.12504>
- 829 Grice, H. P. (1989). *Studies in the Way of Words*. Cambridge: Harvard University
830 Press.
- 831 Happé, F. G. (1994). An advanced test of theory of mind: Understanding of story
832 characters' thoughts and feelings by able autistic, mentally handicapped, and
833 normal children and adults. *Journal of Autism and Developmental Disorders*,
834 *24*(2), 129–154.
- 835 Happé, F. G. E. (1993). Communicative competence and theory of mind in autism: A
836 test of relevance theory. *Cognition*, *48*(2), 101–119.
837 [https://doi.org/10.1016/0010-0277\(93\)90026-R](https://doi.org/10.1016/0010-0277(93)90026-R)
- 838 Harkness, K., Sabbagh, M., Jacobson, J., Chowdrey, N., & Chen, T. (2005).
839 Enhanced accuracy of mental state decoding in dysphoric college students.
840 *Cognition & Emotion*, *19*(7), 999–1025.

841 Ivanko, S. L., & Pexman, P. M. (2003). Context Incongruity and Irony Processing.

842 *Discourse Processes*, 35(3), 241–279.

843 https://doi.org/10.1207/S15326950DP3503_2

844 Jorgensen, J., Miller, G. A., & Sperber, D. (1984). Test of the mention theory of

845 irony. *Journal of Experimental Psychology: General*, 113(1), 112.

846 Katz, A. N., Blasko, D. G., & Kazmerski, V. A. (2004). Saying What You Don't

847 Mean: Social Influences on Sarcastic Language Processing. *Current Directions in*

848 *Psychological Science*, 13(5), 186–189.

849 <https://doi.org/10.1111/j.0963-7214.2004.00304.x>

850 Katz, A. N., & Pexman, P. M. (1997). Interpreting Figurative Statements: Speaker

851 Occupation Can Change Metaphor to Irony. *Metaphor and Symbol*, 12(1), 19–41.

852 https://doi.org/10.1207/s15327868ms1201_3

853 Kidd, D. C., & Castano, E. (2013). Reading literary fiction improves theory of mind.

854 *Science*, 342(6156), 377–380.

855 Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., . . .

856 Rutter, M. (2000). The Autism Diagnostic Observation Schedule—Generic: A

857 standard measure of social and communication deficits associated with the

858 spectrum of autism. *Journal of Autism and Developmental Disorders*, 30(3),

859 205–223.

860 Mar, R. A., Oatley, K., Hirsh, J., Dela Paz, J., & Peterson, J. B. (2006). Bookworms

861 versus nerds: Exposure to fiction versus non-fiction, divergent associations with

862 social ability, and the simulation of fictional social worlds. *Journal of Research in*

863 *Personality*, 40(5), 694–712.

- 864 McDonald, S. (2000). Neuropsychological Studies of Sarcasm. *Metaphor and Symbol*,
865 15(1-2), 85–98. <https://doi.org/10.1080/10926488.2000.9678866>
- 866 Nichols, S., & Stich, S. P. (2003). *Mindreading: An integrated account of pretence,*
867 *self-awareness, and understanding other minds.* Clarendon Press/Oxford
868 University Press.
- 869 Noveck, I. (2018). *Experimental pragmatics: The making of a cognitive science.*
870 Cambridge University Press.
- 871 Oakley, B. F., Brewer, R., Bird, G., & Catmur, C. (2016). Theory of mind is not
872 theory of emotion: A cautionary note on the Reading the Mind in the Eyes Test.
873 *Journal of Abnormal Psychology, 125*(6), 818.
- 874 Olderbak, S., Wilhelm, O., Olaru, G., Geiger, M., Brenneman, M. W., & Roberts, R.
875 D. (2015). A psychometric analysis of the reading the mind in the eyes test:
876 Toward a brief form for research and applied settings. *Frontiers in Psychology, 6*,
877 1503.
- 878 Olkonemi, H., Ranta, H., & Kaakinen, J. K. (2016). Individual differences in the
879 processing of written sarcasm and metaphor: Evidence from eye movements.
880 *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(3),
881 433.
- 882 Peñuelas-Calvo, I., Sareen, A., Sevilla-Llewellyn-Jones, J., & Fern'andez-Berrocal, P.
883 (2019). The “Reading the Mind in the Eyes” test in autism-spectrum disorders
884 comparison with healthy controls: A systematic review and meta-analysis.
885 *Journal of Autism and Developmental Disorders, 49*(3), 1048–1061.
- 886 Pexman, P. M. (2008). It's Fascinating Research: The Cognition of Verbal Irony.
887 *Current Directions in Psychological Science, 17*(4), 286–290.

- 888 <https://doi.org/10.1111/j.1467-8721.2008.00591.x>
- 889 Pexman, P. M., Ferretti, T. R., & Katz, A. N. (2000). Discourse Factors That
890 Influence Online Reading of Metaphor and Irony. *Discourse Processes*, *29*(3),
891 201–222. https://doi.org/10.1207/S15326950dp2903_2
- 892 Pexman, P. M., & Olineck, K. M. (2002). Understanding irony: How do stereotypes
893 cue speaker intent? *Journal of Language and Social Psychology*, *21*(3), 245–274.
- 894 R Core Team. (2020). *R: A language and environment for statistical computing*
895 [Manual]. Vienna, Austria: R Foundation for Statistical Computing.
- 896 Regel, S., Coulson, S., & Gunter, T. C. (2010). The communicative style of a speaker
897 can affect language comprehension? ERP evidence from the comprehension of
898 irony. *Brain Research*, *1311*, 121–135.
899 <https://doi.org/10.1016/j.brainres.2009.10.077>
- 900 Rubio-Fernández, P., Mollica, F., Oraa Ali, M., & Gibson, E. (2019). How do you
901 know that? Automatic belief inferences in passing conversation. *Cognition*, *193*,
902 104011. <https://doi.org/10.1016/j.cognition.2019.104011>
- 903 Schwoebel, J., Dews, S., Winner, E., & Srinivas, K. (2000). Obligatory processing of
904 the literal meaning of ironic utterances: Further evidence. *Metaphor and Symbol*,
905 *15*(1-2), 47–61.
- 906 Searle, J. (1979). Metaphor. In A. Ortony (Ed.), *Metaphor and Thought* (pp.
907 76–116). New York: Cambridge University Press.
908 <https://doi.org/10.1017/CBO9780511609213.006>
- 909 Spaulding, S. (2020). What is mindreading? *Wiley Interdisciplinary Reviews:*
910 *Cognitive Science*, *11*(3), e1523.

- 911 Sperber, D., & Wilson, D. (1981). Irony and the use-mention distinction. *Philosophy*,
912 3, 143–184.
- 913 Sperber, D., & Wilson, D. (2002). Pragmatics, Modularity and Mind-reading. *Mind*
914 & *Language*, 17(1-2), 3–23. <https://doi.org/10.1111/1468-0017.00186>
- 915 Spotorno, N., Koun, E., Prado, J., Van Der Henst, J.-B., & Noveck, I. A. (2012).
916 Neural evidence that utterance-processing entails mentalizing: The case of irony.
917 *NeuroImage*, 63(1), 25–39. <https://doi.org/10.1016/j.neuroimage.2012.06.046>
- 918 Spotorno, N., & Noveck, I. A. (2014). When is irony effortful? *Journal of*
919 *Experimental Psychology: General*, 143(4), 1649–1665.
920 <https://doi.org/10.1037/a0036630>
- 921 Turcan, A., & Filik, R. (2016). An eye-tracking investigation of written sarcasm
922 comprehension: The roles of familiarity and context. *Journal of Experimental*
923 *Psychology: Learning, Memory, and Cognition*, 42(12), 1867–1893.
924 <https://doi.org/10.1037/xlm0000285>
- 925 Turcan, A., & Filik, R. (2017). Investigating sarcasm comprehension using
926 eye-tracking during reading. *Irony in Language Use and Communication*, 1, 255.
- 927 Turcan, A., Howman, H., & Filik, R. (2020). Examining the role of context in
928 written sarcasm comprehension: Evidence from eye-tracking during reading.
929 *Journal of Experimental Psychology: Learning, Memory, and Cognition*.
- 930 Vellante, M., Baron-Cohen, S., Melis, M., Marrone, M., Petretto, D. R., Masala, C.,
931 & Preti, A. (2013). The “Reading the Mind in the Eyes” test: Systematic review
932 of psychometric properties and a validation study in Italy. *Cognitive*
933 *Neuropsychiatry*, 18(4), 326–354.

- 934 Voracek, M., & Dressler, S. G. (2006). Lack of correlation between digit ratio (2D:
935 4D) and Baron-Cohen's "Reading the Mind in the Eyes" test, empathy,
936 systemising, and autism-spectrum quotients in a general population sample.
937 *Personality and Individual Differences*, *41*(8), 1481–1491.
- 938 Wilson, D., & Sperber, D. (1992). On verbal irony. *Lingua*, *87*(1), 53–76.
939 [https://doi.org/10.1016/0024-3841\(92\)90025-E](https://doi.org/10.1016/0024-3841(92)90025-E)
- 940 Wilson, D., & Sperber, D. (2012). *Meaning and Relevance*. Cambridge: Cambridge
941 University Press. <https://doi.org/10.1017/CBO9781139028370>
- 942 Zehr, J., & Schwarz, F. (2018). *PennController for Internet Based Experiments*
943 (*IBEX*). <https://doi.org/10.17605/OSF.IO/MD832>