Irony detection engages the Precuneus and Inferior Frontal Gyrus and relies on the integration of contextual cues and inferential skills

Elizabeth Valles-Capetillo[1,2], Cristian D. Ibarra[1], Magda Giordano[1*]
[1] Departamento de Neurobiología Conductual y Cognitiva
Instituto de Neurobiología, Universidad Nacional Autónoma de México, Boulevard Juriquilla 3001, Querétaro, 76230, México.
[2]Department of Psychology and the Center for Innovative Research in Autism (CIRA), University of Alabama, Tuscaloosa, AL, USA
email: edvallescapetillo@ua.edu, cristian.ibarra@uabc.edu.mx, giordano@unam.mx*
*Corresponding author

Abstract

It has been suggested that irony is one of the most challenging forms of communication; consequently, it can be a valuable indicator of communication abilities (Caillies et al., 2014). An ironic statement transmits the opposite meaning of its literal counterpart (Grice, 1975). The cognitive processes that may support the interpretation of irony include the Theory of Mind (ToM), executive functions (Pexman, 2008), and processing style (Martin and McDonald, 2003). The brain areas associated with irony detection are the medial prefrontal cortex, inferior frontal gyrus (IFG), posterior superior temporal gyrus (pSTG), precuneus, and inferior parietal lobule. This study analyzes the cognitive processes and neural correlates involved in irony detection in Mexican adults. Forty-five participants underwent a cognitive assessment and performed a contextual discrepancy task during fMRI acquisition. The behavioral results showed that detecting nonliteral statements (irony, unrelated, and white lies) requires ToM and verbal abilities. In addition, white lies detection seemed to involve inhibitory control. Ironic statements were the hardest intention to detect, having the lowest percentage of classification and the slowest latency of classification. Irony detection involved brain areas associated with the ToM (i.e., precuneus) and language (i.e., IFG and pSTG), as was expected based on the results of previous studies. Conversely, the detection of literal and unrelated statements recruited motor areas. No differential activation pattern was found for detection of white lies. Finally, a global perceptual processing style predicted the percent change in the BOLD signal in the IFG for all the nonliteral and literal statements.

keywords: irony, fMRI, mentalization, inhibitory control, perceptual processing, contextual discrepancy task

1. Introduction

Language is used to communicate ideas, thoughts, and feelings through the combination of sounds, gestures, and symbols (Balconi, 2010). Nevertheless, understanding the sounds that are part of a language, or the meaning of the words, is not enough to understand the richness of language's meaning in social interactions (Martin & McDonald, 2003). The ability to infer communicative intentions is a key process in human communication (Bambini et al., 2011) because allows the detection of the correct meaning between different possible interpretations of the same statement (Bosco et al., 2017). It has been suggested that irony is one of the most challenging forms of communication (Wilson & Sperber, 1981) it can be a valuable indicator of communication abilities (Caillies et al., 2014). According to the standard pragmatic view, an ironic statement can be defined as one that transmits the opposite meaning of its literal counterpart (Grice, 1975).

The cognitive processes that may support the interpretation of irony include the Theory of Mind (ToM), executive functions (Pexman, 2008), and perceptual processing (Martin & McDonald, 2003). The ToM, or mentalizing, refers to the inferences we make regarding mental states, a relevant ability since it is the mental states that determine actions (Frith & Frith, 2006). The relation between ToM and irony has been reported in behavioral studies in children with cerebral palsy (Caillies et al., 2014). Neuroimaging studies have also reported this relation, which have shown activation in areas associated with ToM (e.g., temporo parietal junction, MPFC, and the precuneus) during irony detection (Spotorno et al., 2012). Executive functions (EF) are general-purpose control mechanisms regulating human cognition and action (Miyake & Friedman, 2012). There are three core EFs, inhibition, working memory, and cognitive flexibility, from which higher-order EFs are built (Diamond, 2013). Inhibition refers to inhibitory control, including self-control (behavioral inhibition) and interference control (selective attention and cognitive inhibition) (Diamond, 2013). The three core EFs have been associated with the detection of irony in older adults (Gaudreau et al., 2015). In typically developing children, inhibitory control has been associated with irony comprehension (Caillies et al., 2014). Additionally, pragmatic language requires context to derive meaning; thus, integrating the utterance with its surrounding context is crucial (Martin & McDonald,

2003). The ability to integrate sources of information has been termed “central coherence,” weak central coherence (WCC) results in attention being given to small pieces at the expense of coherent global patterns of information (Frith, 1989 in Martin and McDonald, 2003). WCC is a bias in a cognitive style that is evident from the perceptual level (visual illusions) to the most complex tasks, such as pragmatic understanding (Booth & Happé, 2010; Martin & McDonald, 2003).

Pragmatic language difficulties can be found in various clinical populations, including traumatic brain injury, right-hemisphere damage, schizophrenia, neurodegenerative disorders, attention-deficit hyperactivity disorder (ADHD), and Autism Spectrum Disorder (ASD) (Booth & Happé, 2010; Caillies et al., 2014; Holtgraves & Giordano, 2017; Martin & McDonald, 2003). Studies in these clinical populations help clarify the role of ToM, EFs, intellectual capacity, and cognitive biases in pragmatic language abilities. It should be noted that among non-clinical individuals, there is variability in those aspects that appear to be affected in the clinical population. One example is central coherence, which varies from strong—focusing on meaning or gestalt—to weak—focusing on details—in the general population (Booth and Happé, 2010). In their study, Boot and Happé (2010) evaluated whether differences in detail focus could be separated from differences in intellectual ability in typically developing children and adolescents (8 -25 years of age; mean age 14.5 years). To measure central coherence, they used a verbal task that required participants to complete sentence stems. In this sample, they found that a significant proportion of the variance in completion scores could not be explained by age, intellectual quotient, or gender. This result could reflect individual differences in cognitive style. Another example is social communication disability, which is proposed to vary between autism and normality (Baron-Cohen, Wheelwright, Skinner, et al., 2001). Baron-Cohen et al. (2001) designed a brief, self-administered questionnaire, the *Autism Spectrum Quotient (AQ),* to evaluate “where any given individual adult, with normal intelligence, lies on this continuum.” The AQ evaluates five areas, i.e., social skills, communication, attention to detail, attention switching, and imagination. Baron-Cohen et al. (2001) found that participants with a diagnosis of Asperger/High Functioning Autism scored significantly higher than matched controls, that within the control group, men scored higher than women, and that there were differences among occupations, replicating earlier findings from the

group. Specifically, the authors found that scientists scored higher than nonscientists; and that “within the sciences, mathematics, physical scientists, computer scientists, and engineers scored higher than the more human or life-centered sciences of medicine (including veterinary science) and biology.”

The brain areas that have been associated with irony detection are the medial prefrontal cortex (MPFC)—either bilaterally or unilaterally—(Shibata et al., 2010; Spotorno et al., 2012; H. Uchiyama et al., 2006; H. T. Uchiyama et al., 2012; Wakusawa et al., 2007; Wang et al., 2006). In the left hemisphere, the inferior frontal gyrus (IFG) (Spotorno et al., 2012), the inferior parietal lobe (IPL) (Rapp et al., 2010; Spotorno et al., 2012), pre-supplementary and motor area (Shibata et al., 2010), precentral gyrus (Rapp et al., 2010), and precuneus (Shibata et al., 2010). And in the right hemisphere, the precentral gyrus (Rapp et al., 2010; Shibata et al., 2010), the dorsolateral prefrontal cortex (dlPFC) (Spotorno et al., 2012), the inferior frontal gyrus (Spotorno et al., 2012), the temporal superior gyrus (Wang et al., 2006) and the cingulate gyrus (Rapp et al., 2010; H. Uchiyama et al., 2006). The results described above come from articles with significant methodological variability concerning the native language of the participants, the modality of the stimuli, and the writing system; none included Spanish-speaking participants. Despite the methodological variability, a meta-analysis showed that the dorsomedial prefrontal cortex (dmPFC), rostromedial prefrontal cortex (rmPFC), inferior frontal gyrus (IFG), and posterior superior temporal gyrus (pSTG) presented consistent activation during irony detection (Reyes-Aguilar et al., 2018).

The present study aimed to analyze the cognitive processes involved in irony detection, including social cognition (e.g., ToM), EF, perceptual processing, and verbal abilities, and find the neural correlates in a sample of neurotypical Mexican adults. Participants from different professional categories (i.e., science, humanities, and engineering) were included to increase the variability in the sample. We expected to replicate and expand the findings of the meta-analysis (Reyes-Aguilar et al., 2018) in terms of the brain regions involved in irony detection and to improve on the previous studies by including a thorough cognitive evaluation of the participants.

2. Method

2.1. Participants

Participants were asked to fill out a general data form with information about their level of education, sex, and age. The experiment was completed by 45 participants (22 female) with a mean age of 26.69 ± 5.83. Of the 45 participants, 15 were studying or had a degree related to science (7 female), 12 non-scientific disciplines (e.g., humanities, 6 female), and 18 in engineering (9 female). Participants reported no history of psychiatric or neurological illness, were right-handed and signed a written informed consent to participate in the study. Handedness was measured with the Edinburgh handedness inventory (Oldfield, 1971). The experimental protocol was approved by the Ethical Committee of the Neurobiology Institute of the Universidad Nacional Autónoma de México (#047.H.RM).

2.2. Behavioral paradigm

To assess irony detection, a contextual discrepancy task for Mexican adults was created (Valles-Capetillo et al., 2022). Fifty-six social contexts and 14 statements were constructed. Each social context category—i.e., ironic, white lies, literal and unrelated——generated an environment that modified the intention of the statement that followed. Each statement was used with all four categories of social contexts. The task was created in Psychopy (3.0.1 version) (Peirce, 2007), and the stimuli were presented in the auditory modality without prosodic changes (see Figure 1). The participants heard the social context, then the statement, and in the last screen, a question regarding the classification of the statement was presented in text modality, with four options.

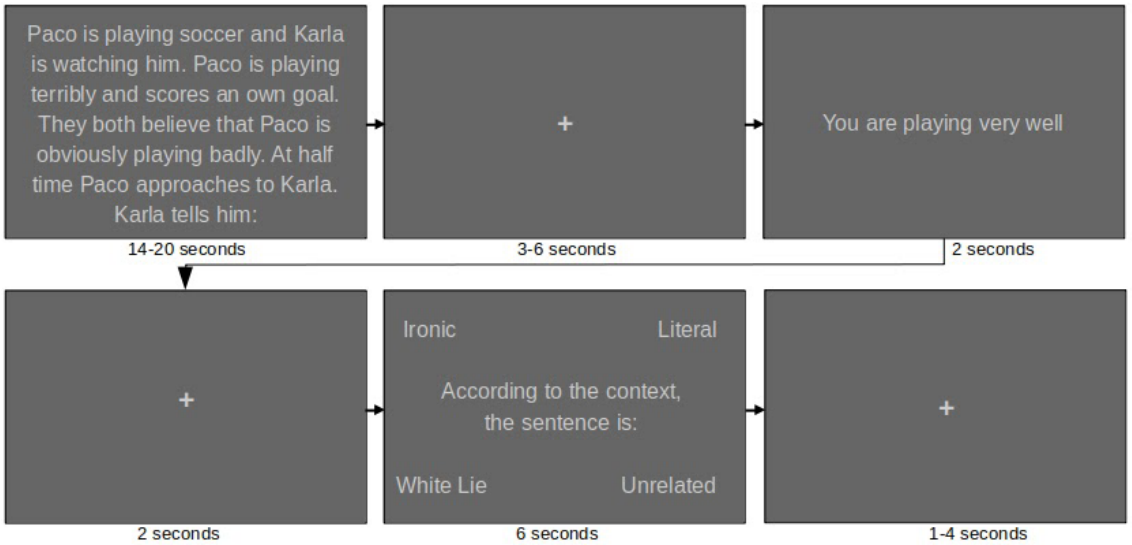

Figure 1.- Contextual discrepancy task. The context and statement were presented in audio without changes in prosody (note: text was not presented). The classification screen was presented in text modality. Stimuli were presented in semi-random order (Valles-Capetillo et al., 2022).

2.3. Cognitive tests

A battery of tests was used to evaluate the cognitive processes associated with irony detection. The cognitive variables of interest were social cognition, executive functions, perceptual processing, and verbal abilities. For social cognition tests, we used tests that evaluate ToM, empathy, use of sarcasm, and social communication abilities. ToM was assessed with the Short Story Task (SST) (Dodell-Feder et al., 2013) and the Reading the Mind in the Eyes Test (RMET) (Baron-Cohen, Wheelwright, Hill, et al., 2001). Empathy was assessed with the Interpersonal Reactivity Index (IRI), which has four scales: fantasy, perspective-taking, empathic concern, and personal distress (Fernández et al., 2011). To assess the degree to which an individual is likely to use sarcasm, the sarcasm self-report scale was used (SSS) (Ivanko et al., 2004). The Autism Spectrum Quotient (AQ) was selected to assess the degree of social abilities, communication, attention to detail, attention switching, and imagination; this inventory has proved to be sensitive in non-clinical populations as described before (Baron-Cohen, Wheelwright, Hill, et al., 2001).

For executive functions, the Psychology Experiment Building Language software (PEBL) was used (Mueller & Piper, 2014). The EF evaluated were updating, assessed with the n-back test; planning, assessed with the Tower of London; inhibitory control, assessed with the go-no-go test; and for working memory, digit span test. Perceptual processing was assessed with the Perceptual Reasoning Scale of WAIS-IV (Wechsler, D., 2008) and the local-global test (Gerlach & Poirel, 2018; Navon, 1977) using Psychtoolbox-3 (Brainard, 1997). The local-global task has been used to evaluate weak central coherence in children with neurodevelopmental disorders, such as ASD and Williams syndrome (Bernardino et al., 2012). Verbal abilities were assessed with the verbal comprehension index (VCI) from the Wechsler Intelligence Scale for Adults, fourth edition (WAIS-IV) (Weschler, 2008); and verbal fluency, assessed with the BANFE (*Batería Neuropsicológica de Funciones Ejecutivas*) (Flores et al., 2008).

A Shapiro test showed that the results of the behavioral task were not normally distributed ($p < 0.01$). Levene's test showed unequal variances between the statement categories ($p < 0.01$). Therefore, non-parametric statistics were used. The percentages and standard deviation are presented for the classification accuracy and the median and interquartile range (IQR) for the latency of classification. A comparison between statement categories was performed using a Friedman test for repeated measures and a Durbin-Conover pairwise comparison analysis with Bonferroni correction.

To find out which were the best cognitive predictors of behavioral performance during the contextual discrepancy task, we used a stepwise regression model for accuracy and latency of classification using R. The stepwise regression starts with no predictors, then sequentially adds the most contributive predictors, then removes the variables that do not improve the model fit (Bruce et al., 2020). The predictors were selected by category of the cognitive process evaluated: social cognition, executive functions, perceptual processing, and verbal abilities. We included the total score for the RMET, SST, IRI, SSS, and AQ for social cognition. In addition, we included reaction time (rt) for the RMET. For EF, we included the go-no-go, Tower of London, digit span, and n-back. For perceptual processing, we included the local global mean accuracy for switching and non-switching, the difference in the

rt between switching and non-switching, and the PRI. And for verbal abilities, verbal fluency and VCI were selected.

2.4. Functional Magnetic Resonance Imaging

2.4.1 MRI data acquisition

The MRI session was performed on a 3-Tesla GE MR750 scanner with a 32-channel head coil. The acquisition parameters were, for T1w: TE of 0.003 seconds, TR of 0.008, flip angle of 12°, and a 256 x 256 matrix; for the functional sequence: TE of 0.04, TR of 2 seconds, 90° flip angle, and 64 x 64 matrix.

2.4.3 Preprocessing

The dicom to nifti format conversion was performed using dcm2niix v1.0.20190902 (X. Li et al., 2016). The preprocessing of the MRI data was performed using the fmriprep pipeline (Esteban et al., 2019).

2.4.4 Task-fMRI procedure

The contextual discrepancy task presentation was optimized for the MRI scanner session using OPTSEQ (Dale et al., 1999) and divided into three runs. Each run had a duration of 10 minutes. We asked the participants to infer the speaker's intention when they heard the statement and to use the information provided by the context to do this task. For the participants to become familiar with the task, before starting the task within the scanner, a training task was given in which they classified 4 stimuli (one per category). The stimuli of the training task were different from those used in the task inside the scanner. NordicNeuroLab goggles (Bergen, Norway) were used for task presentation and a response device (Cedrus Lumina, California, USA), both compatible with the scanner.

2.4.5. Whole brain analysis

Subsequent analyzes were performed with the FMRIB Software Library (FSL; v. 6.0.5) (Jenkinson et al., 2012). For the first (run), second (participant), and third (group) level analyses, a general linear model (GLM) analysis was performed using the FEAT tool (Woolrich et al., 2004). In the first level analysis, the BOLD signal was analyzed during the presentation of the event of interest (i.e., statements). To analyze the events, a GLM was performed using 15 regressors, four for each

statement category (i.e., irony, literal, unrelated, and white lies), one for fixation cross, and 11 for physiological regressors. The physiological regressors were cerebrospinal fluid, white matter, global signal, dvars, frame wise displacement, and 6 motion regressors (translation and rotation of the x, y, z coordinates). Because fmriprep does not perform smoothing, a 6 mm smoothing was applied at this level, and the significance threshold was adjusted at the cluster level ($z = 2.3$ and $p < 0.05$). To analyze the differences between the statement categories, 11 contrasts were analyzed: 4 contrasts comparing each statement category with the fixation cross (e.g., irony > fixation cross), 4 comparing each statement category with the other three (e.g., irony > white lies + literal + no relation), and 3 comparing irony with other statement categories (e.g., irony > white lies).

For the second level analysis (subject level), a fixed effects GLM was performed. And for the third-level analysis (group level), a mixed-effects GLM (i.e., Flame 1) was performed. For both analyses, the threshold parameters of the first level analysis were maintained with cluster correction ($z = 2.3$, $p < 0.05$). The functional activation areas (i.e., labels) were identified using the corresponding MNI152 template coordinates. The labels were obtained using the MNI <-> Talairach tool with the Brodmann areas in the Bioimage Suite Web (Papademetris et al., 2006). This tool defined the Brodmann areas in MNI space (Lacadie et al., 2008).

2.4.5 Percentage change in BOLD signal

To analyze the relationship between the activation in the brain during comprehension of the statements and the cognitive processes we measured, the percentage of BOLD signal was extracted, per subject and condition, from regions previously associated with irony using the feat query tool. To obtain the percent of BOLD signal we selected the contrast that compared each statement category (ironic, white lies, literal and unrelated) with the fixation cross (e.g., ironic statements > fixation cross). The ROIs were selected from a meta-analysis that studied irony (Reyes-Aguilar et al., 2018); they were the dorsomedial PFC ($x= 2$, $y = 54$, $z= 20$), IFG ($x= -50$, $y = 16$, $z= 4$), pSTG ($x= -60$, $y = -40$, $z= 12$), and rostromedial PFC ($x= 2$, $y = 56$, $z= 16$). The ROIs had a radius of 6 mm and were constructed using the fslmaths function.

To find out the best cognitive predictors of the BOLD signal during the presentation of each statement category, we used a stepwise regression model using R. The predictors were selected by category of the cognitive process evaluated: social cognition, executive functions, perceptual processing, verbal abilities, as described regarding the behavioral performance in the contextual discrepancy task.

3. Results

3.1. Behavioral

3.1.1. Contextual discrepancy task

The ironic statements showed the lowest percentage of correct answers (71.43 ± 28.16), followed by white lies (78.57 ± 20.04), and literal (85.71 ± 13.52), while unrelated statements showed the highest accuracy (92.86 ± 14.64). Ironic statements required the longest latency of classification (seconds) (Md = 1.91, IQR = 1.01), followed by literal (Md = 1.72, IQR = 0.87), white lies (Md = 1.49, IQR = 1.20), and unrelated statements (Md = 1.31, IQR = 0.84).

The results show that there are significant differences in the percentage of correct answers ($X^2_{\text{Friedman}}(3) = 45.68$, $p < 0.001$) and latency of classification ($X^2_{\text{Friedman}}(3) = 31.81$, $p < 0.001$). In the percentage of correct answers, significant differences were found between the statement categories of irony compared with literal ($p < 0.05$), unrelated ($p < 0.001$), and white lies ($p < 0.001$); also, between the unrelated statements compared with literal ($p < 0.001$) and white lies ($p < 0.001$; see Figure 2A). latency of classification showed significant differences between the ironic statement compared to unrelated ($p < 0.001$) and white lies ($p < 0.001$), and between literal compared with unrelated ($p < 0.001$) and white lies ($p < 0.001$; see Figure 2B).

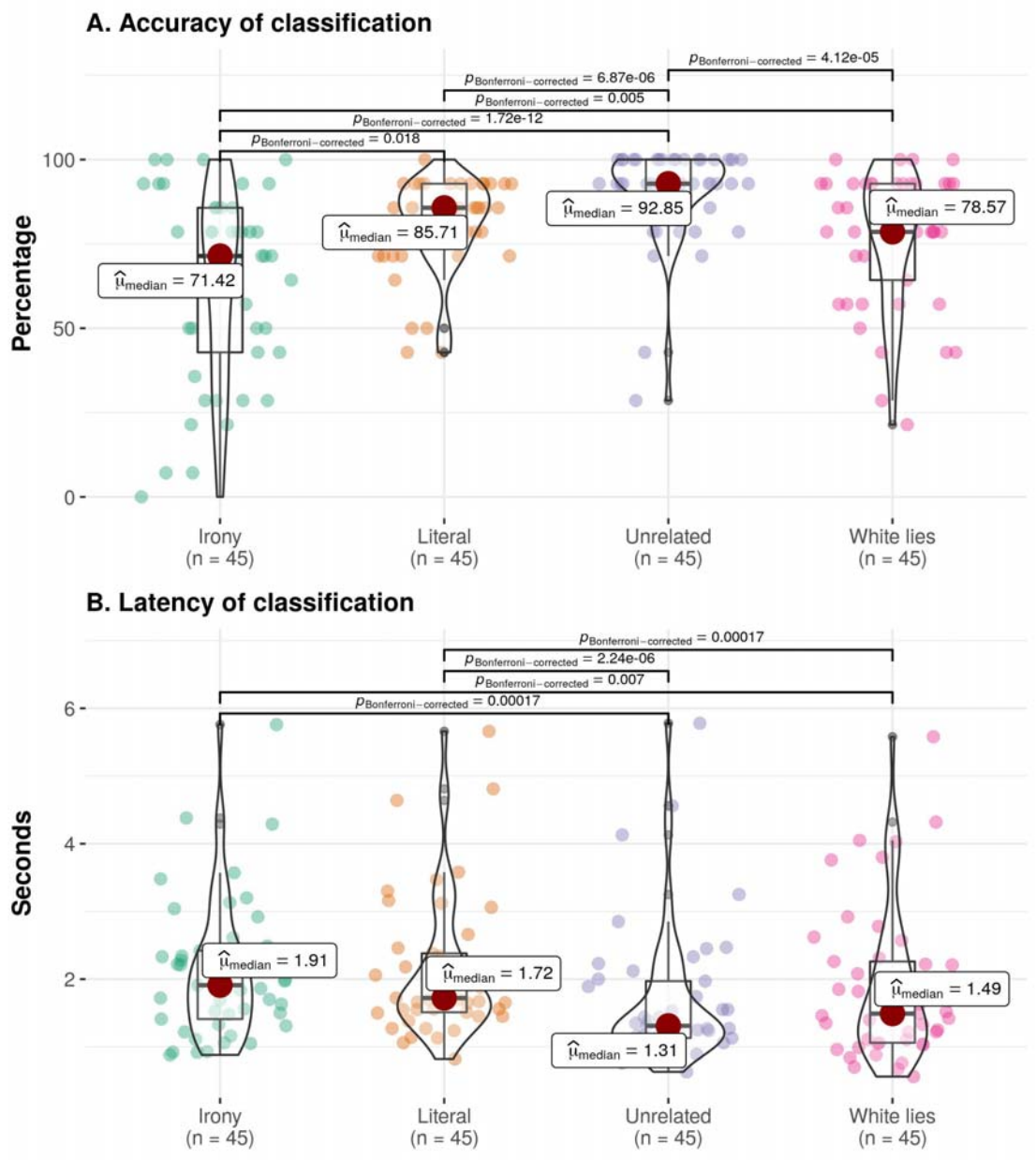

Figure 2.- Percentage of correct answers and latency of classification for each statement category. **A**. Ironic statements resulted in a significantly lower percentage of correct answers, followed by white lies, and literal, while unrelated, had the highest percentage of correct answers. **B**. Latency of classification showed significant differences between the statement categories of irony compared with unrelated and white lies and between literal and unrelated and white lies. The plots show the density curves, and the box plots show the median (red circle), mean (thick line), interquartile range (rectangle), and the lower/upper adjacent values (black lines

stretched from the rectangle), and scatter plot. The significant differences between statement categories are reported with the p-value.

3.1.2. Contextual discrepancy task and cognitive processes

As described in the methods, we used a stepwise regression model to find the best cognitive predictors of behavioral performance during the contextual discrepancy task. The predictors were selected by category of the cognitive process evaluated: social cognition, executive functions, verbal abilities, and perceptual processing (Table 1 and Supplementary Table 1).

Accuracy of irony detection was predicted by scores in AQ ($t = -1.55$, $p = 0.13$) and RMET ($t = 2.588$, $p < 0.05$); in combination, they made a significant contribution ($F(2,41) = 4.16$, $p < 0.05$, adjusted-R2 = 0.12). Also, by VCI ($t = 2.83$, $p < 0.01$; $F(1,41) = 8.01$, $p < 0.01$, adjusted-R2 = 0.14). Similarly, the accuracy of unrelated statements detection was predicted by scores in AQ ($t = -2.37$, $p < 0.05$) and RMET ($t = 4.54$, $p < 0.0001$); in combination, they made a significant contribution ($F(2,41) = 12.1$, $p < 0.0001$, adjusted-R2 = 0.34). VCI ($t = 3.41$, $p < 0.01$) also predicted the accuracy of unrelated statements detection ($F(1,41) = 3.42$, $p < 0.01$, adjusted-R2 = 0.19). Accuracy of white lies detection was predicted by scores in RMET ($t = 1.27$, $p < 0.05$; $F(1,42) = 4.64$, $p < 0.05$, adjusted-R2 = 0.08), scores in the go-no-go task ($t = 2.02$, $p < 0.05$) ($F(1,42) = 4.10$, $p < 0.05$, adjusted-R2 = 0.07), and VCI ($t = 3.45$, $p < 0.01$; $F(1,41) = 11.93$, $p < 0.01$, adjusted-R2 = 0.20). For accuracy of literal statements detection, there were no significant predictors (Table 1, Figures 3A and 3B, and Supplementary Table 2).

For latency of classification, scores in AQ ($t = 3.27$, $p < 0.01$) and SST ($t = 3.30$, $p < 0.01$) predicted latency of classification of ironic statements; in combination, they made a significant contribution ($F(2,41) = 6.83$, $p < 0.001$, adjusted-$R2 = 0.21$). Scores in AQ ($t = 2.19$, $p < 0.05$) and SST ($t = 1.61$, $p = 0.11$) predicted latency of classification of unrelated statements; in combination, they made a significant contribution ($F(2,41) = 3.13$, $p < 0.05$, adjusted-$R2 = 0.09$). Similarly, scores in AQ ($t = 2.82$, $p < 0.01$) and SST ($t = 2.16$, $p < 0.05$) predicted latency of classification of white lies; in combination, they made a significant contribution. Scores in the go-no-go task ($t = -2.09$, $p < 0.05$) also predicted latency of classification of white lies ($F(1,41) = 4.37$, $p < 0.05$, adjusted-$R2 = 0.07$). For literal statements, there were no significant predictors (see Table 1 and Supplementary Table 2).

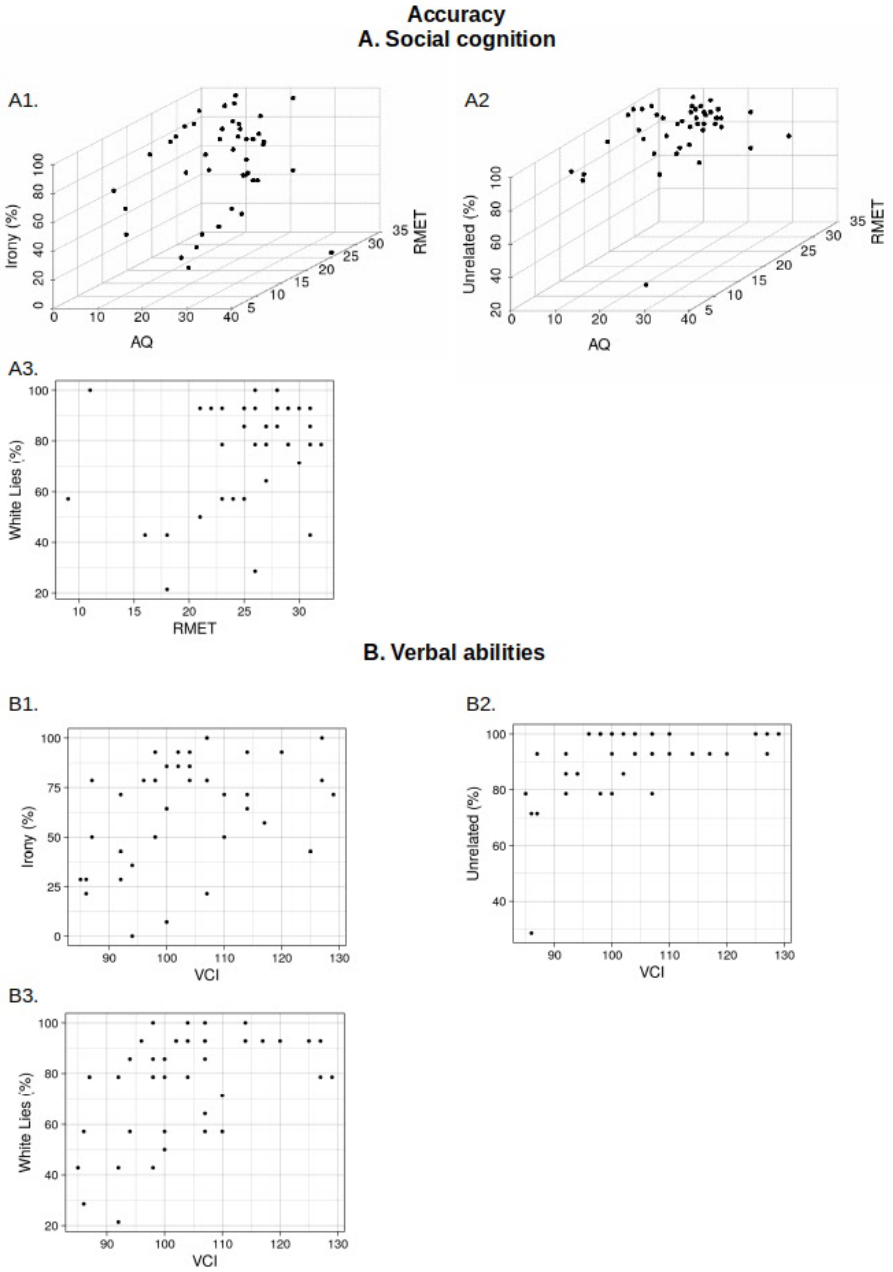

Figure 3. Figures show the relationship between the cognitive predictors' scores and the behavioral results of the contextual discrepancy task (percent accuracy). Panel A shows the role of social cognitive processes (AQ and RMET). Panel B showed the role of verbal abilities, i.e., VCI. For statistical results see: Table 1, Supplementary Table 2, and section 3.1.2.

AQ: autism quotient; RMET: reading the mind in the eyes task, VCI: verbal comprehension index from the WAIS.

Table 1. Stepwise regressors for behavioral results, by cognitive processes and statement categories

Model / dependent variable	Social cognition	Executive functions	Perceptual processing	Verbal abilities
Irony				
Accuracy	0.12*	0.03	n.s.	0.14**
Latency of classification	0.21**	0.04	0.03	n.s.
Literal				
Accuracy	n.s.	n.s.	0.02	n.s.
Latency of classification	0.083	n.s.	0.02	0.027
Unrelated				
Accuracy	0.34***	n.s.	n.s.	0.19***
Latency of classification	0.09*	n.s.	n.s.	n.s.
White lies				
Accuracy	0.078*	0.067*	n.s.	0.20***
Latency of classification	0.16**	0.07*	0.02	n.s.

The adjusted-R2 are presented for each model. For details on the statistical results see Supplementary Table 2.
p = * < 0.05, ** < 0.01, *** < 0.001.

n.s. = no significant predictors survived in the stepwise regression

3.2. Functional Magnetic Reasoning Imaging

3.2.1. The contrast between one statement category versus the others

The contrast irony > literal+unrelated+white lies showed significant differential activation in the precuneus, in Brodmann's area 31 (BA) (see Figure 4A; Table 2). The contrast literal > irony+unrelated+white lies showed significant differential activation in three clusters: the right premotor area (BA 6); the left primary motor area (BA 4), extending to the primary sensory area (BA 1); and the right putamen, extending to the caudate (see Figure 4B; Table 2). The contrast unrelated > irony+literal+white lies showed significant differential activation in three clusters: the left premotor area (BA 6), extending to the anterior cingulate gyrus (BA 24); the left visual-motor area (BA 7), extending to the anterior supramarginal gyrus (BA 40); and the right primary motor area (BA 4), extending to the premotor area (BA 6) (see Figure 4C; Table 2). The contrast white lies > irony+literal+unrelated did not show significant results.

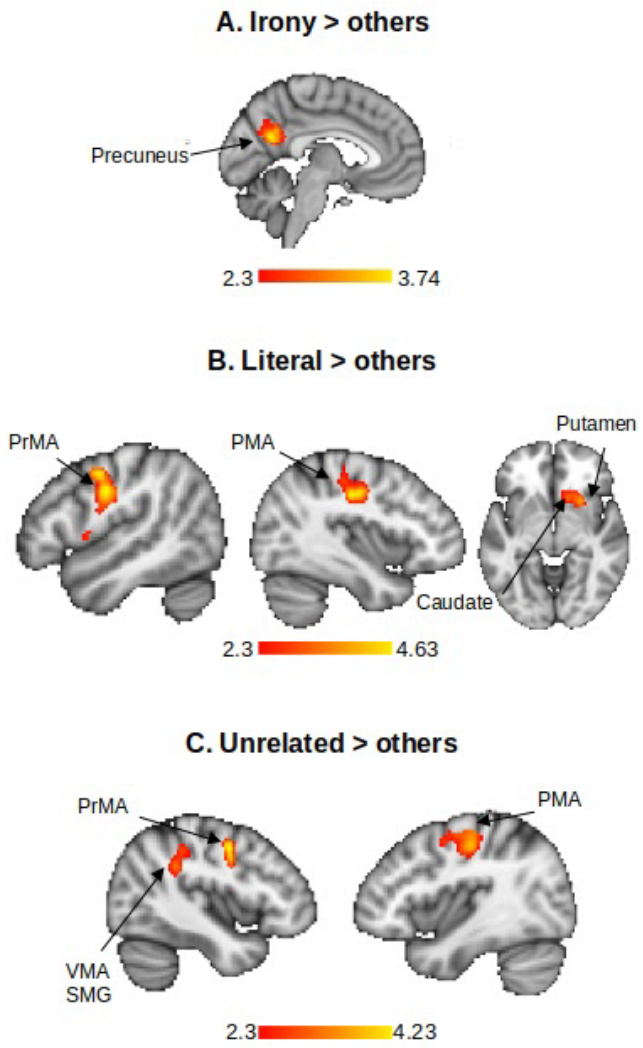

Figure 4.- Results of the contrast between each statement category and the others. Panel A shows the contrast irony > literal+unrelated+white lies; this contrast shows significant differential activation in the precuneus. Panel B shows the contrast literal > irony+unrelated+white lies. Panel C shows the contrast of unrelated > irony+unrelated+white lies. The contrast white lies > irony+literal+unrelated did not show significant differential activation. Maps were thresholded at $z < 2.3$, $p < 0.05$. The cluster extension Montreal Neurological Institute (MNI) coordinates are listed in Table 2. PrMA = premotor area, PMA = primary motor area, VMA = ventral motor area, SMG = supramarginal gyrus.

3.2.2. The contrast between ironic statements and other statement categories

The irony > literal contrast showed significant differential activation in three clusters, the precuneus, extending into the left motor visual area (BA 7) and the right visual-motor area (BA 7); the left angular gyrus (BA 39), extending into the medial temporal gyrus (BA 31) and posterior superior temporal gyrus (BA 22); and the right IFG (pars triangularis, BA 45,), extending into the dorsolateral prefrontal cortex (BA 9) and IFG (pars opercularis, BA 44) (see Figure 5A; Table 3 for MNI coordinates).

The irony > unrelated contrast showed significant differential activation in two clusters, the right cuneal cortex (BA 18), extending into the associative visual area (BA 18) and lingual gyrus (BA 17) and secondary visual cortex (BA 18); the precuneus (BA 31), extending into the ventral cingulate gyrus (BA 23) and the lingual gyrus (BA 18) (see Figure 5B; Table 3 for MNI coordinates).

The irony > white lies contrast showed significant differential activation in two clusters, the right supramarginal gyrus (BA 40), extending into the posterior superior temporal gyrus (BA 22), precuneus (BA 31), cingulate gyrus (BA 31), and the IFG (pars opercularis, BA 44); and the left primary sensory cortex (BA 1,), extending into BA 5 of the primary association cortex. (see Figure 5C; Table 3 for MNI coordinates).

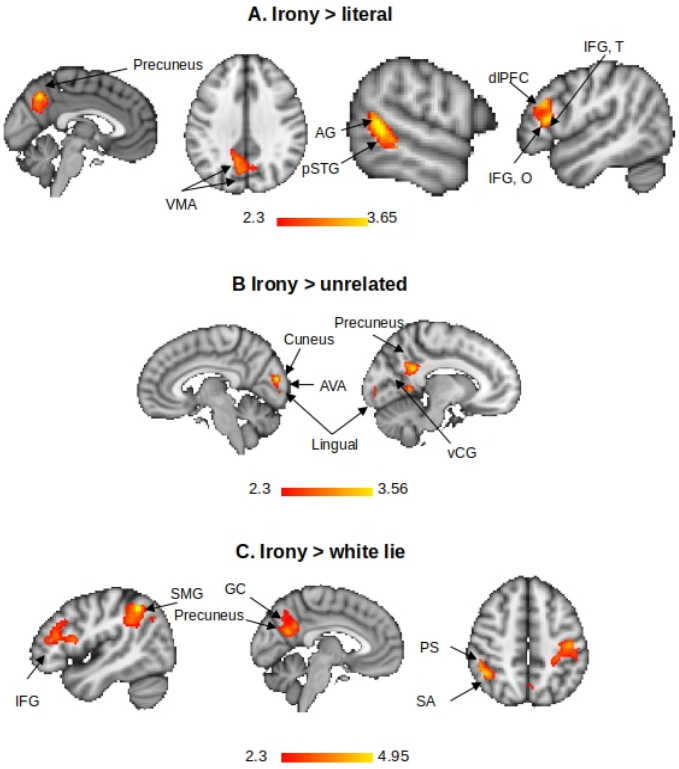

Figure 5.- Contrasts of irony versus other statement categories. Panel A shows the contrast irony > literal; Panel B, irony > unrelated; panel C, irony > white lies. Maps were thresholded at $z < 2.3$, $p < 0.05$. The cluster extension MNI coordinates are listed in Table 3. VMA = visual motor area, AG = angular gyrus, pSTG = posterior superior temporal gyrus, dlPFC = dorsolateral prefrontal cortex, IFG T = inferior frontal gyrus triangularis, IFG O = inferior frontal gyrus opercularis, AVA = associative visual area, VCG = ventral cingulate gyrus, SMG = supramarginal gyrus, GC = cingulate gyrus, SA = primary association cortex, PS = primary sensory cortex.

Table 2. Brain areas exhibiting significant activation in whole brain analysis for contrasts between one statement category versus the others, according to GLM analysis. The x, y, and z coordinates are in MNI space; labels were obtained using the MNI <-> Talairach tool with the Brodmann areas in the Bioimage Suite Web (Papademetris et al., 2006). LH: left hemisphere, RH: right hemisphere. Results were corrected at the cluster level.

Contrast	Vox	Hem	Region	Cluster center - MNI Coordinate			Z	Individual foci - MNI Coordinate			Label
				x	y	z		x	y	z	
Irony > All	576	LH	Precuneous cortex (BA 31)	-4	-60	24	3.74				
Literal > All	1433	RH	Premotor + Supplementary motor (BA 6)	50	-6	32	4.63				
	660	LH	Primary motor (BA 4)	-38	-20	38	4.51				
							4.38	-42	-12	40	LH Premotor + Supplementary motor (BA 6)
							2.91	-40	-26	48	LH Primary sensory (BA 1)
	623		Putamen	26	16	-6	3.74				
							3.26	38	-6	12	RH Primary motor (BA 4)
							3.22	14	20	-4	RH Caudate
Unrelated > All	963	LH	Premotor + Supplementary motor (BA 6)	-42	-4	44	4.23				
							3.19	-14	-8	58	LH Superior frontal gyrus

							3.14	-2	-4	48	Ventral anterior cingulate (BA 24)
	789	LH	Visual motor (BA 7)	-32	-46	42	4.29				
							3.32	-62	-32	32	LH Supramarginal gyrus (BA 40)
							3.17	-42	-40	30	LH Parietal opercularis cortex
	446	RH	Primary motor (BA 4)	40	-18	46	3.73				
							2.76	40	0	48	RH Premotor + Supplementary motor (BA 6)
White lies > All			n.s								

Table 3. Brain areas exhibiting significant activation in whole brain analysis for contrasts between ironic statements and the other categories, according to GLM analysis. The x, y, and z coordinates are in MNI space, labels were obtained using the MNI <-> Talairach tool with the Brodmann areas in the Bioimage Suite Web (Papademetris et al., 2006).. LH: left hemisphere, RH: right hemisphere. Results were corrected at the cluster level.

Contrast	Vox	Hem	Region	Cluster center - MNI Coordinate			Z	Individual foci - MNI Coordinate			Label
				x	y	z		x	y	z	
Irony > White lies	3014	RH	Supramarginal gyrus (BA 40)	46	-50	46	4.95				
							3.92	-4	-62	22	LH Precuneous
							3.79	-6	-54	30	LH Cingulate gyrus
							3.72	44	-42	32	RH Supramarginal gyrus (BA 40)
							3.6	46	14	14	RH Inferior frontal gyrus, opercularis (BA 44)
	768	LH	Postcentral gyrus (Primary sensory, BA 1)	-42	-20	48	4.02				
							3.53	-30	-34	48	
Irony > Literal	742	LH	Precuneous cortex (Visual motor, BA 7)	-2	-62	44	3.65				
							3.34	-10	-60	26	LH Precuneous (BA 31)
							3.09	-6	-68	34	LH Visual motor (BA 7)
							3.02	8	-64	36	RH visual motor (BA 7)
	615	LH	Angular gyrus (BA 39)	-58	-50	10	3.68				
							3.63	-58	-48	6	LH Middle temporal gyrus (BA 21)

	535	RH	Inferior frontal gyrus, pars triangularis (BA 45)	50	24	14	3.48				
							3.36	50	28	26	RH Dorsolateral prefrontal cortex, BA 9)
							3.07	40	18	16	RH Inferior frontal gyrus, opercularis (BA 44)
							2.75	46	36	16	RH Dorsolateral prefrontal cortex (BA 46)
Irony > Unrelated	553	RH	Secundary visual (BA 18)	6	-84	16	3.54				
							3.2	22	-90	12	RH Secondary visual (BA 18)
							3.17	2	-88	0	RH Primary visual (BA 17)
							2.99	-6	-92	-2	LH Secondary visual (BA 18)
							2.9	16	-74	0	RH Secondary visual (BA 18)
	462	LH	Precuneous (BA 31)	-2	-60	24	3.77				
							3.44	-8	-50	24	LH Ventral cingulate (BA 23)
							3.1	-10	-52	2	LH Secondary visual (BA 18)

3.2.3. Percentage of BOLD signal and cognitive processes

As described in the methods, to analyze the relationship between the activation in the brain during the presentation of the statements and the cognitive processes we measured, the percentage change in the BOLD signal was extracted from regions previously associated with irony. The selected ROI were the dorsomedial PFC, IFG, pSTG, and rostromedial PFC, and the signal was extracted from the contrast that compared each statement category (ironic, white lies, literal and unrelated) with the fixation cross.

Scores on social cognition tests predicted BOLD signal change in the IFG during literal ($F(2,41) = 6.14$, $p < 0.01$, adjusted-R2 = 0.19), unrelated ($F(2,41) = 4.13$, $p < 0.05$, adjusted-R2 = 0.07) and white lies ($F(2,41) = 9.14$, $p < 0.01$, adjusted-R2 = 0.27) detection. For the detection of literal statements, the predictors were RMET scores ($t = 2.04$, $p < 0.05$) and RMET rt ($t = 3.22$, $p < 0.01$), for unrelated statements, RMET rt ($t = 2.03$, $p < 0.05$), and for white lies, IRI ($t = 1.67$, $p = 0.10$) and RMET rt ($t = 3.95$, $p < 0.001$) (Table 4, and Supplementary Table 3).

Table 4. Stepwise regressors for the percentage of BOLD signal change associated with detecting statement categories in Regions of Interest (ROI) by cognitive process.

Model / dependent variable	Social cognition	Executive functions	Processing style	Verbal abilities
Irony				
LH-dmPFC	0.088	0.038	0.036	n.s.

LH-rmPFC	0.086	0.067	0.053	n.s.
LH-STG	0.037	n.s.	n.s.	n.s.
LH-IFG	n.s.	0.037	0.202**	n.s.
Literal				
LH-dmPFC	n.s.	0.031	n.s.	0.041
LH-rmPFC	n.s.	0.043	n.s.	0.042
LH-STG	n.s.	0.064	n.s.	n.s.
LH-IFG	0.193**	0.035	0.328***	n.s.
Unrelated				
LH-dmPFC	n.s.	0.073*	n.s.	n.s.
LH-rmPFC	n.s.	0.06	n.s.	n.s.
LH-STG	n.s.	0.067*	n.s.	n.s.
LH-IFG	0.068*	n.s	0.171**	0.023
White lies				

LH-dmPFC	n.s.	n.s.	n.s.	n.s.
LH-rmPFC	0.073	n.s.	n.s.	n.s.
LH-STG	n.s.	0.058	n.s.	n.s.
LH-IFG	0.275***	0.032	0.25**	n.s.

The adjusted-R2 are presented for each model. For details on the statistical analysis see Supplementary Table 3.
p = * < 0.05, ** < 0.01, *** < 0.001.
n.s. = no significant predictors survived in the stepwise regression

Executive functions predicted BOLD signal in the dmPFC during unrelated ($F(1,42) = 4.37$, $p < 0.05$, adjusted-R2 = 0.073) and STG detection ($F(1,42) = 4.09$, $p < 0.05$, adjusted-R2 = 0.067). The significant predictor for dmPFC ($t = 2.09$, $p < 0.05$) and STG ($t = 2.02$, $p < 0.05$) was go-no-go.

Perceptual processing also predicted BOLD signal change in the IFG during detection of ironic ($F(2,41) = 6.45$, $p < 0.01$, adjusted-R2 = 0.20), literal ($F(3,40) = 7.99$, $p < 0.001$, adjusted-R2 = 0.32), unrelated $F(2,41) = 5.43$, $p < 0.01$, adjusted-R2 = 0.17), and white lies $F(3,40) = 5.94$, $p < 0.01$, adjusted-R2 = 0.25) statements. The predictors for the detection of ironic statements were the local-global rt difference ($t = 3.42$, $p < 0.01$) and the no-switch accuracy ($t = -3.09$, $p < 0.01$). The predictors for the detection of literal statements were the PRI ($t = 2.26$, $p < 0.05$), local-global mean rt difference ($t = 4.5$, $p <0.001$), and no switching accuracy ($t = -3.76$, $p < 0.001$). The predictors for the detection of unrelated statements were local-global mean rt difference ($t = 3.2$, $p <0.01$) and no switching accuracy ($t = -2.3$, $p < 0.05$). The predictors for the detection of white lies were the PRI ($t = 2.40$, $p < 0.05$), local-global mean reaction time difference ($t = 3.61$, $p <0.001$),

and no switching accuracy ($t = -3.30$, $p < 0.01$). No other test scores were significant predictors of BOLD signal change in any other ROI.

4. Discussion

This paper aimed to evaluate the cognitive processes and the neural areas associated with detecting ironic statements. The contextual discrepancy task we employed showed that ironic statements were difficult to detect, having the lowest accuracy and highest classification latency. This result agrees with our previous behavioral study (Valles-Capetillo et al., 2022) and with the proposal that irony is one of the most difficult communication forms (Wilson & Sperber, 1981). Regarding the neural regions involved, we found that the precuneus, which has been associated with ToM, presented consistent activation during the detection of the ironic statements. This activation was shown in the contrast of irony to all the other statements, the contrast of irony to literal statements, and unrelated statements. These results are consistent with those of Shibata et al. (2010) and Varga et al. (2013), who studied irony comprehension in Japanese, and Hungarian participants, respectively.

To discover which cognitive processes could be involved in detecting irony, we used a battery of tests that evaluated the domains of social cognition, executive functions, perceptual processing, and verbal abilities. These processes have been previously proposed as pertinent (Martin & McDonald, 2003; Pexman, 2008) and found to be relevant in experimental studies in clinical and non-clinical populations (Booth & Happé, 2010; Caillies et al., 2014; Holtgraves & Giordano, 2017). In the social cognition domain, as was expected, ToM predicted the accuracy and latency of the classification of the ironic statements. The other statement categories, unrelated and white lies, were also predicted by ToM; this relation was not seen for the literal statements. Perhaps ToM facilitates the detection of nonliteral statements, especially because their intended meaning is not immediately apparent since it varies according to the context (Bohrn et al., 2012; Lisofsky et al., 2014; Pexman, 2008; Shamay-Tsoory et al., 2005; Spotorno et al., 2012). Interestingly, the AQ questionnaire predicted the nonliteral statements' classification

latency. This questionnaire evaluates abilities associated with autism, such as the degree of social abilities, communication, attention to detail, attention switching, and imagination (Baron-Cohen, Wheelwright, Skinner, et al., 2001). We found that participants with better abilities, as evaluated by the AQ, took less time to detect nonliteral statements. Together these results show that social cognitive processes play a fundamental role in understanding linguistic information according to the context.

Regarding executive functions, inhibitory control was a predictor for the accuracy and latency of the classification of white lies. The fact that white lies are rated as less sincere than other statements (Valles-Capetillo et al., 2022) suggests that inhibition of a preponderant interpretation plays an important role (Christ et al., 2009).

Concerning verbal abilities, the VCI was a positive predictor for the detection of ironic, unrelated, and white lies statements. These non-literal statements bear indirect messages and thus require substantial linguistic abilities (Grice, 1975). Indeed, neuroimaging studies have shown that nonliteral language recruits areas associated with language and their homologous such as the right IFG (Bohrn et al., 2012; Reyes-Aguilar et al., 2018).

As for the results of the whole-brain neuroimaging analysis, the contrast between irony and all other statements, as well as between irony and other statement categories, showed significant differential activation in the left precuneus, a region that has been associated with ToM, and is a key region of the default mode network that “simultaneously interacts with both the default-mode and frontoparietal networks to distinguish distinct cognitive states” (Utevsky et al., 2014). The precuneus seems to be involved in mental imagery and the construction of inferential meanings (Amodio & Frith, 2006; Cavanna & Trimble, 2006; Frith & Frith, 2006; Gallagher et al., 2000; Mar et al., 2008; Schnell et al., 2016; Schurz et al., 2014). This brain region also showed differential activation during speech act recognition, consistent with its role in inferential processes (Licea-Haquet et al., 2021). It is worth noting that the precuneus did not show differential activation during

the detection of other statement categories. Instead, the contrast of the literal and unrelated statements versus the other categories showed differential activation in areas associated with motor activation (e.g., BA 4 and 6). The contrast of white lies versus all other statement categories did not show a differential activation pattern.

We expected that the detection of irony would show significant activation in those areas found by a previous meta-analysis (Reyes-Aguilar et al., 2018), i.e., the MPFC, dorsomedial, and rostromedial (BA 9); IFG (BA 44), and STG (BA 22). Indeed, we found differential activation in the right IFG in the contrasts between irony versus literal statements and white lies. This region has been associated with processes related to language, such as the detection of unusual or incongruous semantic fields (Jung-Beeman, 2005; Obert et al., 2016), detection of semantic violations, reanalysis of meanings (Bašnáková et al., 2014), the integration of linguistic information into context (Shtyrov & Pulvermüller, 2007), and the selection of the best interpretation according to the social context (Egorova et al., 2016). Event-related potentials (ERPs) studies have shown an increase in the N400–a potential associated with detecting anomalous linguistic information–when evaluating irony on the basis of prosody or context (Cornejo et al., 2007; Regel et al., 2010; Regel & Gunter, 2017).

The contrast between ironic and literal statements also showed activation in the pSTG. This area has been associated with processes related to social communication, such as ToM (Amodio & Frith, 2006; Gallagher et al., 2000; Mar et al., 2008), action observation (Yang et al., 2015), extracting social information about a situation (Egorova et al., 2016; Redcay, 2008) and the integration of paralinguistic cues with contextual information (Egorova et al., 2016). ERPs studies of irony employing a variety of contexts and sentences have shown an increase in the P600, which has been associated with integrating information to facilitate irony detection (Regel et al., 2010; Spotorno et al., 2013; Weissman & Tanner, 2018). While eye-tracking studies using context or stories and statements have reported an increment in regressions during the reading of ironic stimuli, the regression has been associated with the integration of necessary information to infer

irony (Filik et al., 2017; Filik & Moxey, 2010; Olkoniemi, Johander, et al., 2019; Olkoniemi, Strömberg, et al., 2019). Together, these findings underscore the relevance of information integration for irony detection.

In addition, the contrast of ironic versus literal statements showed differential activation in the right dlPFC. This region has been associated with the processes of goal-oriented attention, task switching, planning, novelty seeking (Jones & Graff-Radford, 2021), problem-solving, and working memory (Jones & Graff-Radford, 2021; Mansouri et al., 2007). This region has also been associated with those aspects of complex language processing, such as ambiguity, nonliteral meaning and when cues need to be integrated to infer the linguistic meaning (Hertrich et al., 2021).

Compared to the neuroimaging results obtained in the contrast between irony and the other statement categories, the contrasts between literal and unrelated versus all the other categories (e.g., literal > irony + unrelated + white lies), showed differential activation in the primary motor and premotor areas. The primary motor cortex (BA 4) has been associated with observing motor tasks, integrating multiple sensory inputs, inhibiting involuntary movement, and imagining motor sequences (Bhattacharjee et al., 2021). The premotor cortex (BA 6) has been associated with the observation of goal-directed actions (Binkofski & Buccino, 2006; Chouinard & Paus, 2006; Morin & Grèzes, 2008), which allows recognition of actions carried out by others; the mirror neuron system, involved in social interactions and attribution of intentions to others (Bhattacharjee et al., 2021; Morin & Grèzes, 2008), and recognition of multimodal actions (Binkofski & Buccino, 2006). The differential activation of these regions during the detection of literal and unrelated statements versus the other categories suggests that their detection requires additional resources, such as inhibiting involuntary movement and imagining motor sequences.

One of our hypotheses was that a global perceptual processing style–central coherence–that focuses on broad patterns of information would be associated with detecting irony. Indeed, the perceptual processing scores predicted the percent change in

the BOLD signal in the IFG during ironic, literal, unrelated, and white lies statements detection; the local-global test was the significant predictor. Although it did not significantly predict behavioral performance on the contextual discrepancy task (i.e., accuracy and latency of classification). This result agrees with previous findings associating the IFG with sentence-context integration (Shtyrov & Pulvermüller, 2007). Also, it could reflect an essential role of a global perceptual processing style in detecting linguistic information during social interactions.

Taken together, these behavioral and neuroimaging results agree with Pexman's (2008) constraint satisfaction model in that irony detection is predicted by ToM and verbal abilities. In terms of psycholinguistic properties relevant to the detection of irony, Attardo (2000) proposes that ironic statements are relevant —have relation to the context—, are inappropriate to the context, and are meant to convey the intended meaning to the listener. To detect irony, then, a listener needs to compare the statement to the contextual cues present and to infer the speaker's intention. In the contextual discrepancy task, the context varied to give the statement its intended meaning while the statements remained the same. The neuroimaging contrasts between irony and the other statement categories consistently showed the activation of the precuneus, suggesting that this region is involved in the comparison between context and statement and the inference about the speaker's intention. Intriguingly, the supramarginal gyrus showed differential activation in the contrast of irony versus white lies but not in the opposite contrast. This region appears to be involved in deception (Meier et al., 2021). These two statement categories are similar in that they give relevant but inappropriate information in relation to the context but differ in terms of the speaker's intention. Indeed, the contrast between white lies and the other categories showed no differential pattern of brain activation, suggesting that the less sincere intention of the speaker, which is the essential element in white lies, was not distinctive enough.

5. Conclusion

In conclusion, selecting a heterogeneous sample of participants allowed us to discern the effect of different cognitive processes on irony detection. The behavioral results (accuracy and classification latency) showed that detecting nonliteral statements (irony, unrelated, and white lies) requires ToM and verbal abilities. In addition, white lies detection seemed to involve inhibitory control. Regarding the brain regions associated with irony detection, differential activation patterns were found in areas associated with the ToM (i.e., precuneus) and language (i.e., IFG and pSTG), as was expected on the basis of a previous meta-analysis of pragmatic language (Reyes-Aguilar et al., 2018). Conversely, the detection of literal and unrelated statements recruited motor areas (e.g., primary motor area, premotor area, visual-motor area). No differential activation was found for white lies, possibly because their detection engages processes similar to those for the other categories, with the exception of the deceitful intention. Finally, a global perceptual processing style predicted the percent change in the BOLD signal in the IFG for all the nonliteral and literal statements. For future studies, it would be useful to evaluate the effect of psycholinguistic properties during irony detection. For example, how ironic, sincere, appropriate, or relevant a statement was perceived and its relation to changes in BOLD signal. Also, including participants with difficulties in understanding social communication, such as ASD, would help to increase the sample variability and improve our understanding of the relation between linguistic, behavioral, cognitive, and brain correlates of irony detection. Finally, our stimuli involve contextual discrepancy for their detection, but other cues are present during social communication, such as facial expression and prosody. Using other stimuli to present the statements, such as videos, would help to study irony detection more ecologically.

Declaration of interest

Conflict of interest: none

Acknowledgements

Diola Elizabeth Valles Capetillo is a doctoral student from the Programa de Doctorado en Ciencias Biomédicas, Universidad Nacional Autónoma de México (UNAM) and has received CONACyT fellowship 755580. This study was supported by grants from DGAPA-PAPIIT (IN 203818) and CONACyT (Fronteras No. 225-2015 and scholarship to E. V. C. No. 755580). We greatly acknowledge their experimental guidance to Luis Concha Noyola and Juan Fernandez Ruiz. For technical support, we thank Erick Pasaye and Leopoldo González, the technical staff, and the laboratory Lavis (UNAM-Juriquilla). For experimental support, we want to thank Abril Dominguez. For their careful review of the manuscript, we want to thank Jessica González Norris and Mike Jeziorski.

Author Contributions

Elizabeth Valles-Capetillo: designed research, performed research, contributed unpublished reagents/analytic tools, analyzed data and wrote the paper. **Cristian Ibarra**: performed research, contributed unpublished reagents/analytic tools, and wrote the paper. **Magda Giordano**: designed research, contributed unpublished reagents/analytic tools, wrote the paper, methodology, funding acquisition, project administration and supervision.

References

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, *7*(4), 268–277. https://doi.org/10.1038/nrn1884

Balconi, M. (2010). Biological Basis of Linguistic and Communicative Systems: From Neurolinguistics to Neuropragmatics. In M. Balconi (Ed.), *Neuropsychology of Communication* (pp. 3–27). Springer Milan. https://doi.org/10.1007/978-88-470-1584-5_1

Bambini, V., Gentili, C., Ricciardi, E., Bertinetto, P. M., & Pietrini, P. (2011). Decomposing metaphor processing at the cognitive and neural level through functional magnetic

resonance imaging. *Brain Research Bulletin*, *86*(3–4), 203–216. https://doi.org/10.1016/j.brainresbull.2011.07.015

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. https://doi.org/10.1017/S0021963001006643

Baron-Cohen, Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, *31*(1), 5–17. https://doi.org/10.1023/a:1005653411471

Bašnáková, J., Weber, K., Petersson, K. M., Van Berkum, J., & Hagoort, P. (2014). Beyond the language given: The neural correlates of inferring speaker meaning. *Cerebral Cortex*, *24*(10), 2572–2578. https://doi.org/10.1093/cercor/bht112

Bhattacharjee, S., Kashyap, R., Abualait, T., Annabel Chen, S.-H., Yoo, W.-K., & Bashir, S. (2021). The Role of Primary Motor Cortex: More Than Movement Execution. *Journal of Motor Behavior*, *53*(2), 258–274. https://doi.org/10.1080/00222895.2020.1738992

Binkofski, F., & Buccino, G. (2006). The role of ventral premotor cortex in action execution and action understanding. *Journal of Physiology-Paris*, *99*(4), 396–405. https://doi.org/10.1016/j.jphysparis.2006.03.005

Bohrn, I. C., Altmann, U., & Jacobs, A. M. (2012). Looking at the brains behind figurative language--a quantitative meta-analysis of neuroimaging studies on metaphor, idiom, and irony processing. *Neuropsychologia*, *50*(11), 2669–2683. https://doi.org/10.1016/j.neuropsychologia.2012.07.021

Booth, R., & Happé, F. (2010). “Hunting with a knife and ... fork”: examining central coherence in autism, attention deficit/hyperactivity disorder, and typical development with a linguistic

task. *Journal of Experimental Child Psychology*, *107*(4), 377–393. https://doi.org/10.1016/j.jecp.2010.06.003

Bosco, F. M., Parola, A., Valentini, M. C., & Morese, R. (2017). Neural correlates underlying the comprehension of deceitful and ironic communicative intentions. *Cortex*, *94*, 73–86. https://doi.org/10.1016/j.cortex.2017.06.010

Bruce, P., Bruce, A., & Gedeck, P. (2020). *Practical Statistics for Data Scientist* (Second Edition). O'Reilly Media, Inc. https://www.oreilly.com/library/view/practical-statistics-for/9781491952955/

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*(11), 3191–3198. https://doi.org/10.1016/j.ridd.2014.08.002

Cavanna, A. E., & Trimble, M. R. (2006). The precuneus: a review of its functional anatomy and behavioural correlates. *Brain*, *129*(3), 564–583. https://doi.org/10.1093/brain/awl004

Chouinard, P. A., & Paus, T. (2006). The Primary Motor and Premotor Areas of the Human Cerebral Cortex. *The Neuroscientist*, *12*(2), 143–152. https://doi.org/10.1177/1073858405284255

Christ, S. E., Van Essen, D. C., Watson, J. M., Brubaker, L. E., & McDermott, K. B. (2009). The Contributions of Prefrontal Cortex and Executive Control to Deception: Evidence from Activation Likelihood Estimate Meta-analyses. *Cerebral Cortex*, *19*(7), 1557–1566. https://doi.org/10.1093/cercor/bhn189

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical Surface-Based Analysis. *NeuroImage*, *9*(2), 179–194. https://doi.org/10.1006/nimg.1998.0395

Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, *64*(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Dodell-Feder, D., Lincoln, S. H., Coulson, J. P., & Hooker, C. I. (2013). Using fiction to assess

mental state understanding: a new task for assessing theory of mind in adults. *PloS One*, *8*(11), e81279. https://doi.org/10.1371/journal.pone.0081279

Egorova, N., Shtyrov, Y., & Pulvermüller, F. (2016). Brain basis of communicative actions in language. *NeuroImage*, *125*, 857–867. https://doi.org/10.1016/j.neuroimage.2015.10.055

Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., Kent, J. D., Goncalves, M., DuPre, E., Snyder, M., Oya, H., Ghosh, S. S., Wright, J., Durnez, J., Poldrack, R. A., & Gorgolewski, K. J. (2019). fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nature Methods*, *16*(1), 111–116. https://doi.org/10.1038/s41592-018-0235-4

Fernández, A. M., Dufey, M., & Kramp, U. (2011). Testing the psychometric properties of the Interpersonal Reactivity Index (IRI) in Chile: Empathy in a different cultural context. *European Journal of Psychological Assessment*, *27*(3), 179–185. https://doi.org/10.1027/1015-5759/a000065

Filik, R., Brightman, E., Gathercole, C., & Leuthold, H. (2017). The emotional impact of verbal irony: Eye-tracking evidence for a two-stage process. *Journal of Memory and Language*, *93*, 193–202. https://doi.org/10.1016/j.jml.2016.09.006

Filik, R., & Moxey, L. M. (2010). The on-line processing of written irony. *Cognition*, *116*(3), 421–436. https://doi.org/10.1016/j.cognition.2010.06.005

Flores, J., Ostrosky, F., & Lozano, A. (2008). Batería de Funciones Frontales y Ejecutivas: Presentación. *Revista Neuropsicología, Neuropsiquiatría y Neurociencias*, *8*(1), 141–158.

Frith, C. D., & Frith, U. (2006). The Neural Basis of Mentalizing. *Neuron*, *50*(4), 531–534. https://doi.org/10.1016/J.NEURON.2006.05.001

Gallagher, H., Happe, F., Brunswick, N., Fletcher, P., Frith, U., & Frith, C. (2000). Gallagher HL, Happé F, Brunswick N, Fletcher PC, Frith U, Frith CD. Reading the mind in cartoons and stories: an fMRI study of ‘theory of mind’ in verbal and nonverbal tasks. Neuropsychologia

38: 11-21. *Neuropsychologia*, *38*, 11–21. https://doi.org/10.1016/S0028-3932(99)00053-6

Gaudreau, G., Monetta, L., Macoir, J., Poulin, S., Laforce, R. J., & Hudon, C. (2015). Mental State Inferences Abilities Contribution to Verbal Irony Comprehension in Older Adults with Mild Cognitive Impairment. *Behavioural Neurology*, *2015*, e685613. https://doi.org/10.1155/2015/685613

Grice, H. P. (1975). Logic and Conversation. *Speech Acts*, 41–58. https://doi.org/10.1163/9789004368811_003

Hertrich, I., Dietrich, S., Blum, C., & Ackermann, H. (2021). The Role of the Dorsolateral Prefrontal Cortex for Speech and Language Processing. *Frontiers in Human Neuroscience*, *15*, 645209. https://doi.org/10.3389/fnhum.2021.645209

Holtgraves, T., & Giordano, M. (2017). Parkinson's Disease Without Dementia. In *Perspectives in Pragmatics, Philosophy and Psychology* (pp. 379–407). https://doi.org/10.1007/978-3-319-47489-2_15

Ivanko, S. L., Pexman, P. M., & Olineck, K. M. (2004). How Sarcastic Are You?: Individual Differences and Verbal Irony. *Journal of Language and Social Psychology*, *23*(3), 244–271. https://doi.org/10.1177/0261927X04266809

Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., & Smith, S. M. (2012). FSL. *NeuroImage*, *62*(2), 782–790. https://doi.org/10.1016/j.neuroimage.2011.09.015

Jones, D. T., & Graff-Radford, J. (2021). Executive Dysfunction and the Prefrontal Cortex. *Continuum (Minneapolis, Minn.)*, *27*(6), 1586–1601. https://doi.org/10.1212/CON.0000000000001009

Jung-Beeman, M. (2005). Bilateral brain processes for comprehending natural language. *Trends in Cognitive Sciences*, *9*(11), 512–518. https://doi.org/10.1016/j.tics.2005.09.009

Lacadie, C., Fulbright, R. K., Arora, J., Constable, R., & Papademetris, X. (2008). Brodmann Areas defined in MNI space using a new Tracing Tool in BioImage Suite. *Proceedings of the 14th*

Annual Meeting of the Organization for Human Brain Mapping, *771*.

Licea-Haquet, G. L., Reyes-Aguilar, A., Alcauter, S., & Giordano, M. (2021). The Neural Substrate of Speech Act Recognition. *Neuroscience*, *471*, 102–114. https://doi.org/10.1016/j.neuroscience.2021.07.020

Lisofsky, N., Kazzer, P., Heekeren, H. R., & Prehn, K. (2014). Investigating socio-cognitive processes in deception: A quantitative meta-analysis of neuroimaging studies. *Neuropsychologia*, *61*, 113–122. https://doi.org/10.1016/j.neuropsychologia.2014.06.001

Mansouri, F. A., Buckley, M. J., & Tanaka, K. (2007). Mnemonic Function of the Dorsolateral Prefrontal Cortex in Conflict-Induced Behavioral Adjustment. *Science*, *318*(5852), 987–990. https://doi.org/10.1126/science.1146384

Mar, R., Djikic, M., & Oatley, K. (2008). Effects of reading on knowledge, social abilities, and selfhood: Theory and empirical studies. *Directions in Empirical Literary Studies*, 127–137.

Martin, I., & McDonald, S. (2003). Weak coherence, no theory of mind, or executive dysfunction? Solving the puzzle of pragmatic language disorders. *Brain and Language*, *85*(3), 451–466. https://doi.org/10.1016/S0093-934X(03)00070-1

McGrath, H., Zaveri, H. P., Collins, E., Jafar, T., Chishti, O., Obaid, S., Ksendzovsky, A., Wu, K., Papademetris, X., & Spencer, D. D. (2022). High-resolution cortical parcellation based on conserved brain landmarks for localization of multimodal data to the nearest centimeter. *Scientific Reports*, *12*(1), 18778. https://doi.org/10.1038/s41598-022-21543-3

Meier, S. K., Ray, K. L., Mastan, J. C., Salvage, S. R., & Robin, D. A. (2021). Meta-analytic connectivity modelling of deception-related brain regions. *PloS One*, *16*(8), e0248909. https://doi.org/10.1371/journal.pone.0248909

Miyake, A., & Friedman, N. P. (2012). The Nature and Organization of Individual Differences in Executive Functions: Four General Conclusions. *Current Directions in Psychological Science*, *21*(1), 8–14. https://doi.org/10.1177/0963721411429458

Morin, O., & Grèzes, J. (2008). What is “mirror” in the premotor cortex? A review. *Neurophysiologie Clinique = Clinical Neurophysiology*, *38*(3), 189–195. https://doi.org/10.1016/j.neucli.2008.02.005

Mueller, S. T., & Piper, B. J. (2014). The Psychology Experiment Building Language (PEBL) and PEBL Test Battery. *Journal of Neuroscience Methods*, *222*, 250–259. https://doi.org/10.1016/j.jneumeth.2013.10.024

Obert, A., Gierski, F., Calmus, A., Flucher, A., Portefaix, C., Pierot, L., Kaladjian, A., & Caillies, S. (2016). Neural Correlates of Contrast and Humor: Processing Common Features of Verbal Irony. *PLOS ONE*, *11*(11), e0166704. https://doi.org/10.1371/journal.pone.0166704

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Olkoniemi, H., Johander, E., & Kaakinen, J. K. (2019). The role of look-backs in the processing of written sarcasm. *Memory & Cognition*, *47*(1), 87–105. https://doi.org/10.3758/s13421-018-0852-2

Olkoniemi, H., Strömberg, V., & Kaakinen, J. K. (2019). The ability to recognise emotions predicts the time-course of sarcasm processing: Evidence from eye movements. *Quarterly Journal of Experimental Psychology (2006)*, *72*(5), 1212–1223. https://doi.org/10.1177/1747021818807864

Papademetris, X., Jackowski, M. P., Rajeevan, N., DiStasio, M., Okuda, H., Constable, R. T., & Staib, L. H. (2006). BioImage Suite: An integrated medical image analysis suite: An update. *The Insight Journal*, *2006*, 209.

Peirce, J. W. (2007). PsychoPy--Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1–2), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Pexman, P. M. (2008). It’s Fascinating Research. *Current Directions in Psychological Science*, *17*(4), 286–290. https://doi.org/10.1111/j.1467-8721.2008.00591.x

Rapp, A. M., Mutschler, D. E., Wild, B., Erb, M., Lengsfeld, I., Saur, R., & Grodd, W. (2010). Neural correlates of irony comprehension: The role of schizotypal personality traits. *Brain and Language*, *113*(1), 1–12. https://doi.org/10.1016/j.bandl.2009.11.007

Redcay, E. (2008). The superior temporal sulcus performs a common function for social and speech perception: Implications for the emergence of autism. *Neuroscience & Biobehavioral Reviews*, *32*(1), 123–142. https://doi.org/10.1016/j.neubiorev.2007.06.004

Regel, S., Coulson, S., & Gunter, T. C. (2010). The communicative style of a speaker can affect language comprehension? ERP evidence from the comprehension of irony. *Brain Research*, *1311*, 121–135. https://doi.org/10.1016/j.brainres.2009.10.077

Reyes-Aguilar, A., Valles-Capetillo, E., & Giordano, M. (2018). A Quantitative Meta-analysis of Neuroimaging Studies of Pragmatic Language Comprehension: In Search of a Universal Neural Substrate. *Neuroscience*, *395*, 60–88. https://doi.org/10.1016/j.neuroscience.2018.10.043

Schnell, Z., Varga, E., Tényi, T., Simon, M., Hajnal, A., Járai, R., & Herold, R. (2016). Neuropragmatics and irony processing in schizophrenia – Possible neural correlates of the meta-module of pragmatic meaning construction. *Journal of Pragmatics*, *92*, 74–99. https://doi.org/10.1016/J.PRAGMA.2015.11.004

Schurz, M., Radua, J., Aichhorn, M., Richlan, F., & Perner, J. (2014). Fractionating theory of mind: A meta-analysis of functional brain imaging studies. *Neuroscience & Biobehavioral Reviews*, *42*, 9–34. https://doi.org/10.1016/J.NEUBIOREV.2014.01.009

Shamay-Tsoory, S. G., Tomer, R., & Aharon-Peretz, J. (2005). The Neuroanatomical Basis of Understanding Sarcasm and Its Relationship to Social Cognition. *Neuropsychology*, *19*(3), 288–300. https://doi.org/10.1037/0894-4105.19.3.288

Shibata, M., Toyomura, A., Itoh, H., & ichi Abe, J. (2010). Neural substrates of irony comprehension: A functional MRI study. *Brain Research*, *1308*, 114–123.

https://doi.org/10.1016/j.brainres.2009.10.030

Shtyrov, Y., & Pulvermüller, F. (2007). Early MEG Activation Dynamics in the Left Temporal and Inferior Frontal Cortex Reflect Semantic Context Integration. *Journal of Cognitive Neuroscience*, *19*(10), 1633–1642. https://doi.org/10.1162/jocn.2007.19.10.1633

Spotorno, N., Cheylus, A., Van Der Henst, J.-B., & Noveck, I. A. (2013). What's behind a P600? Integration operations during irony processing. *PloS One*, *8*(6), e66839. https://doi.org/10.1371/journal.pone.0066839

Spotorno, N., Koun, E., Prado, J., Van Der Henst, J. B., & Noveck, I. A. (2012). Neural evidence that utterance-processing entails mentalizing: The case of irony. *NeuroImage*, *63*(1), 25–39. https://doi.org/10.1016/j.neuroimage.2012.06.046

Uchiyama, H., Seki, A., Kageyama, H., Saito, D. N., Koeda, T., Ohno, K., & Sadato, N. (2006). Neural substrates of sarcasm: A functional magnetic-resonance imaging study. *Brain Research*, *1124*(1), 100–110. https://doi.org/10.1016/j.brainres.2006.09.088

Uchiyama, H. T., Saito, D. N., Tanabe, H. C., Harada, T., Seki, A., Ohno, K., Koeda, T., & Sadato, N. (2012). Distinction between the literal and intended meanings of sentences: A functional magnetic resonance imaging study of metaphor and sarcasm. *Cortex*, *48*(5), 563–583. https://doi.org/10.1016/j.cortex.2011.01.004

Utevsky, A. V., Smith, D. V., & Huettel, S. A. (2014). Precuneus is a functional core of the default-mode network. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *34*(3), 932–940. https://doi.org/10.1523/JNEUROSCI.4227-13.2014

Valles-Capetillo, E., Ibarra, C., Martinez, D., & Giordano, M. (2022). A novel task to evaluate irony comprehension and its essential elements in Spanish speakers. *Frontiers in Psychology*, *13*, 963666. https://doi.org/10.3389/fpsyg.2022.963666

Wakusawa, K., Sugiura, M., Sassa, Y., Jeong, H., Horie, K., Sato, S., Yokoyama, H., Tsuchiya, S., Inuma, K., & Kawashima, R. (2007). Comprehension of implicit meanings in social

situations involving irony: A functional MRI study. *NeuroImage*, *37*(4), 1417–1426. https://doi.org/10.1016/j.neuroimage.2007.06.013

Wang, A. T., Lee, S. S., Sigman, M., & Dapretto, M. (2006). Developmental changes in the neural basis of interpreting communicative intent. *Social Cognitive and Affective Neuroscience*, *1*(2), 107–121. https://doi.org/10.1093/scan/nsl018

Weissman, B., & Tanner, D. (2018). A strong wink between verbal and emoji-based irony: How the brain processes ironic emojis during language comprehension. *PLOS ONE*, *13*(8), e0201727. https://doi.org/10.1371/journal.pone.0201727

Weschler, D. (2008). *Wechsler Adult Intelligence Scale | Fourth Edition*. https://www.pearsonassessments.com/store/usassessments/en/Store/Professional-Assessments/Cognition-%26-Neuro/Wechsler-Adult-Intelligence-Scale-%7C-Fourth-Edition/p/100000392.html

Wilson, D., & Sperber, D. (1981). Irony and the use-mention distinction. *1981*. https://www.academia.edu/17030837/Irony_and_the_use_mention_distinction

Yang, D. Y.-J., Rosenblau, G., Keifer, C., & Pelphrey, K. A. (2015). An integrative neural model of social perception, action observation, and theory of mind. *Neuroscience & Biobehavioral Reviews*, *51*, 263–275. https://doi.org/10.1016/J.NEUBIOREV.2015.01.020