The Heart can Lie: A Preliminary Investigation of the Role of Interoception and Theory of Mind in Deception

Abstract

While a large part of the deception literature focuses on lying detection, the factors contributing to one's ability to lie remain unclear. The present study examined the contribution of Theory of Mind (ToM) and interoception on our ability to lie using a directed lie paradigm with two conditions ("Interrogation" and "Polygraph"), designed to enhance each of the two mechanisms. Given the relatively small sample size (n = 26 x 40 trials), special steps were taken to avoid false positives. Our results suggest that various facets of interoceptive abilities are positively related to the self-rated confidence in one's own lies, especially when under the belief that bodily signals are being monitored (i.e., in the "Polygraph" condition). Beyond providing evidence for the role of the body in lying and raising interesting questions for deception science, these results carry practical implications for criminology and lie detection protocols.

Keywords: Deception; Interoception; Theory of Mind; Poygraph; Lying Ability

Word count: 5004

The Heart can Lie: A Preliminary Investigation of the Role of Interoception and Theory of Mind in Deception

Lying - the intentional attempt at instilling a false belief in others (Sip et al., 2012) - is a prevalent phenomenon carrying potentially important consequences. Interestingly, evidence suggests that the successful detection of a lying attempt depends more on the ability of the liar, than on the performance of the lie detector (Bond Jr & DePaulo, 2008; T. R. Levine et al., 2011; Verigin et al., 2019). However, with most of the deception literature focused on deception detection (Masip, 2017; Sternglanz et al., 2019; Viji et al., 2022), the factors contributing to one's ability to lie remain unclear. Nevertheless, some findings suggest a relationship between the propensity to tell lies, and traits that characterize the socially malevolent profile known as the Dark Triad (Paulhus & Williams, 2002), such as narcissism (Zvi & Elaad, 2018) and psychopathy (Rassin et al., 2023). While often conceptualized to be immoral and unconscionable, lying is ubiquitous in everyday life, and being able to lie skillfully can sometimes facilitate interpersonal relationships, helping us avoid conflict or causing emotional harm to others (E. E. Levine & Lupoli, 2022). In fact, recent research shows that certain forms of deception, such as prosocial lies (i.e., false statements told to benefit others; E. E. Levine and Lupoli (2022)), can increase trust (E. E. Levine & Schweitzer, 2015). Moreover, individuals who told altruistic lies were perceived as more benevolent than those who were honest (E. E. Levine & Schweitzer, 2014).

As deception requires the liar to intentionally manipulate the beliefs of others (Burgoon & Buller, 1994; Sip et al., 2012), a significant line of research has been focused on the role of theory of mind (ToM) in lying ability. ToM refers to the ability to infer that others have mental states, such as beliefs, emotions and intentions, distinct from ourselves (Baron-Cohen, 1997; Lee & Imuta, 2021; Wellman et al., 2001). The ability to tell lies, as well as their complexity, have previously been found to be related to higher ToM abilities (Evans & Lee, 2011; Talwar et al., 2007, 2017). However, studies investigating the link between ToM and deception have predominantly been focused on children and neuroatypical individuals (Beaudoin et al., 2020; Bora & Yener, 2017; Roheger et al., 2022), and its importance in healthy adults remains to be clarified.

Besides paying attention to the person we lie to, gauging whether they believe us, some attention is also directed inwards: monitoring our own body and its reactions (e.g., cardiac activity and its related changes such as blushing), which could be used as cues to infer our real intent. This begs the question of the potential role of interoceptive abilities in deception ability. Broadly defined as one's sensitivity to their own internal signals and bodily states (Chen et al., 2021; Murphy et al., 2019; Weiss et al., 2014), Garfinkel et al. (2015) conceptualize interoception as a three-dimensional construct comprising three distinct facets, namely, interoceptive accuracy - the objective ability to monitor internal bodily signals; interoceptive sensibility - the subjective confidence in one's interoceptive accuracy; and interoceptive awareness - the metacognitive ability to correctly evaluate one's interoceptive ability. Interoception has increasingly been tied to subjective perceptual experiences (Connell et al., 2018; Seth et al., 2012), as well as individual differences in executive functions, emotional processing, and decision-making (Barrett & Simmons, 2015; Murphy et al., 2019; Petzschner et al., 2021).

Although few studies exist that investigate the relationship between interoception and deceptive ability *per se*, previous decision-making studies have demonstrated a negative correlation between interoceptive awareness, a metacognitive dimension of interoception, and one's likelihood to make risky decisions (Dunn et al., 2010; Furman et al., 2013). This is in line with the somatic marker hypothesis, which posits that an accurate evaluation of one's bodily signals facilitates the use of such interoceptive feedback to guide rational decision making (Damasio, 1996). Indeed, Sugawara et al. (2020) further reports that individuals who received interoceptive training were more likely to show higher interoceptive accuracy and make reasoned decisions. Given that deciding to lie generally involves a consideration of the potential costs of getting caught, and hence could also be perceived as risky behavior (Kireev et al., 2013), interoception could be construed to be negatively related to lying ability. However, some studies have instead found heightened interoceptive attention (one's self-focus towards internal bodily signals), to predict immoral behaviour, such as cheating (Ditto et al., 2006; Lenggenhager et al., 2013; Williams et al., 2016). Extending these findings to social cognition, Vabba et al. (2022) further reports individuals with lower interoception told significantly less egoistic lies when the social reputational stakes were high, whereas individuals with higher interoception did not exhibit a significant difference in the number of lies told. Given the scarce research on interoception and deception, more studies are herein needed to clarify these mixed findings.

The aim of the present study was to explore the contribution of ToM and interoception abilities on individuals' deception skills, as indicated by their lying confidence, physiological arousal and response time. To this end, we designed a directed-lying paradigm with 2 conditions differing in the nature of their feedback cues. The *Interrogation* condition was designed to emphasize (and preferentially mobilize) ToM-related mechanisms, whereas the *Polygraph* condition was designed to emphasize interoceptive mechanisms. In particular, we expected lying ability (i.e., higher lie confidence, shorter response time and lower physiological arousal), to be positively predicted by individuals' interoceptive abilities in the *Polygraph* condition, and by ToM skills in the *Interrogation* condition. Consistent with the cognitive load approach outlined in several theories of deception (such as the Four-Factor Theory (Riggio et al., 1987) and Activation-Decision-Construction Model (Walczyk et al., 2014)), as well as previous findings which suggest response time as a reliable cue to deception (Gonzalez-Billandon et al., 2019; Walczyk et al., 2009), we regarded shorter response times as a proxy of better lying ability.

Methods

Participants

30 university students from Singapore were recruited through posters, flyers, and online social media platforms, and rewarded with study credits for their time. Four participants were excluded as their data was not recorded due to technical issues. The final sample consists of 26 participants (Mean age = 20.9, SD = 2.0, range:[18, 25], Sex: 65.4% women, 34.6% men). The heart rate of one participant and response time of one participant were excluded from further analysis due to extreme outlying values. To maximize statistical power, the problematic data from these 2 participants were only excluded from analyses involving those measures; all other data were retained for analyses.

This study was approved by the NTU Institutional Review Board (NTU-IRB-2020-09-007). All participants provided their informed consent prior to participation and were awarded with academic credits upon completion of the study.

Measures

Theory of Mind (ToM). Two measures of ToM and its related constructs were administered. The Yoni Task (Shamay-Tsoory & Aharon-Peretz, 2007) is a behavioral task which assesses first and second-order ToM abilities in both cognitive and affective domains. Participants were presented with the face of a character named "Yoni", surrounded by 4 colored pictures of objects or faces, one in each corner of the screen. In total, each participant completed 101 trials - 49 trials assessing their affective ToM abilities, 37 trials assessing their cognitive ToM abilities and 15 control trials. During each trial, participants were shown a question pertaining the item Yoni is referring to, and asked to make responses based on specific corresponding cues such as the directions of Yoni's eye gaze, facial expressions etc., In the control trials, participants made judgements based on Yoni's physical context (physical ToM). More specifically, in first-order trials, participants were instructed to make inferences about Yoni's mental state with regards to the objects surrounding it (e.g., "Yoni is thinking of…"). In more complex second-order trials, participants had to correctly infer the interaction between Yoni and others' mental states (e.g., Yoni is thinking of the fruit that … wants").

The Basic Empathy Scale (BES, Jolliffe & Farrington, 2006), a 20-item self-report questionnaire measuring two dimensions of empathy, namely Cognitive ($\alpha = 0.83$) and Affective ($\alpha = 0.82$) using a 5-point Likert scale was administered. Although ToM and empathy are regarded as distinct psychological constructs, previous research findings point to them being closely related (Gallant et al., 2020; Sebastian et al., 2012). Specifically, empathy is often thought to be an integral component in the affective dimension of ToM (i.e., the ability to infer what someone else is feeling) (Shamay-Tsoory et al., 2010).

Interoception. To assess participants' interoceptive ability, participants completed a Heartbeat Counting Task (HCT, Schandry, 1981) while having their actual heartbeats recorded. During the HCT task, participants were instructed to count the number of heartbeats over 5 trials with varying time intervals (20s, 25s, 30s, 35s, 40s), the order of which was randomized. Interoceptive accuracy was computed from the difference between the estimated number and the real number of heart beats. Interoceptive sensibility was estimated as the average of the confidence ratings presented at the end of each trial. Interoceptive awareness was indexed by the correlation between the objective accuracy and the subjective confidence.

Given its multidimensional nature, the MAIA-2 (Mehling et al., 2012), a 37-item questionnaire which measures 8 distinct facets of interoception including Noticing (e.g., *I notice when I am uncomfortable in my body*; $\alpha = 0.70$), Not-Distracting (e.g., *I try to ignore pain*; $\alpha = 0.87$), Not-Worrying (e.g., *I can stay calm and not worry when I have feelings of discomfort or pain*; $\alpha = 0.68$), Attention Regulation (e.g., *I can refocus my attention from thinking to sensing my body*; $\alpha = 0.85$), Emotional Awareness (e.g., *I notice how my body changes when I am angry*; $\alpha = 0.75$), Self-Regulation (e.g., *I can use my breath to reduce tension*;

$\alpha = 0.62$), Body Listening (e.g., *I listen to information from my body about my emotional state*; $\alpha = 0.88$), and Trust (e.g., *I trust my body sensations*; $\alpha = 0.89$), using 5-point Likert scales, was also administered.

Deception. Using PsychoPy (Peirce et al., 2019), we implemented a directed-lying task in which participants were instructed to briefly answer 80 questions (taken from the Autobiographical Memory Questionnaire - AMQ, Rubin et al., 2003) pertaining to their personal preferences and subjective experiences, by either lying or telling the truth (depending on whether they see "lie" or "truth" written on the screen). Their goal was to make convincing answers, so that truths would be judged as truths by the receiver, and lies as lies. The nature of the receiver was different depending on the condition: participants were told that for half of the trials, they would have to convince another participant that would be observing them from a separate room (COVID regulations were used as a justification) via a webcam connection (*Interrogation* condition). For the remaining trials, participants were tasked to convince a "lie detection machine" that would be assessing their behaviour through their physiological signals (*Polygraph* condition). In reality, there was no real "receiver" and their answers were not judged externally (the study focused on their subjective ratings and reactions). The two conditions were presented in a counter-balanced order, and each comprised of 40 trials (20 truth; 20 lies).

The sequence of each trial was the same for both conditions. Participants were first instructed to respond to a question shown on a computer screen by verbally lying or telling the truth (for half of the trials in each condition, i.e., n = 20). In addition, as past studies have found associations between lying behaviour and type of question phrasing (Walczyk & Cockrell, 2022), each question was phrased either directly (e.g., "What is your favourite sport?"), or indirectly (e.g., "Is your favourite sport Hockey?") to reduce possible confounding effects. Following a short interval (0.7 - 1.5s) to allow time for response preparation, during which "Connecting. . . " was shown on the screen, a social or bio-feedback cue (for *Interrogation* and *Polygraph* conditions respectively) was displayed for a maximum of 10s or until a response was given. Specifically, participants had to provide their answers verbally, and pressed the space key to signify the end of their response. After another short interval (1.5-2.5s), during which "Disconnecting…" was presented on the screen, participants were asked to rate their confidence in how convincing (i.e., likely to get judged as truthful) they perceived their response to be on a visual analog scale.

In the *Interrogation* condition, participants had to provide their answer while receiving social feedback in the form of a video stream of an examiner, and were informed that the examiner would be evaluating the truthfulness of their responses after observing them through the live video feed. In actuality, the displayed video feed were pre-recorded video clips of a confederate's face (staying still with minimal reactions, and with a medical mask, as the experiment was run during COVID restrictions), and the same video-clips were used with all participants. All participants reported believing that the stream was real and that the examiner was really there during the debriefing.

In the *Polygraph* condition, participants had to provide their answer while receiving bio-feedback in the form of physiological signals (ECG, EDA and Respiration), of which they were informed was live feedback of their own physiological activity. In actuality, we displayed pre-recorded video clips of a confederate's physiological signals. All participants reported believing that their own signals were shown during the debriefing.

Three outcome variables were recorded for each trial of the deception task, namely the participants' confidence ratings that their answers (lies or truths) were convincing, the response time (RT) between the question onset and the participant's key press (indicating the end of their verbal answer), and the change in heart rate associated with the response (within a window of 3.5 s).

Finally, on top of the deception task, we also measured participants' self-reported tendency to lie in their everyday life using the *Lie Scale* (Makowski, Pham, Lau, Raine, et al., 2021), a 16-item questionnaire that assesses 4 dispositional lying dimensions - Ability ($\alpha = 0.92$), Frequency ($\alpha = 0.66$), Negativity ($\alpha = 0.66$), and Contextuality ($\alpha = 0.70$).

Procedure

A within-subjects design was used in the present study, which is comprised of 2 sessions, to investigate the roles interoception and ToM play in lying ability. During session 1, participants answered a brief demographic survey as well as a questionnaire regarding their personal preferences and subjective experiences (Autobiographical Memory Questionnaire, AMQ), followed by a series of psychological scales (i.e., BES, MAIA and Lie scale), which were randomly displayed.

During session 2, performed about one week later, the 3 cognitive-behavioural tasks (i.e., the deception task, HCT and the Yoni task) were administered to participants while their physiological signals (including cardiac activity - ECG, respiration -RSP, and electrodermal activity - EDA) were being recorded. The physiological recording devices were set up as follows: ECG was recorded with 3 electrodes placed according to a modified Lead II configuration (Takuma et al., 1995), and respiration was measured using a respiration belt. All signals were recorded at 1000Hz via the BioPac MP160 system (BioPac Systems Inc., USA).

For all participants, session 2 began with the deception task, followed by the Yoni task and the HCT, with the latter two presented in a randomized order. In the directed-lying task, items of the AMQ were presented as stimuli, with participants' recorded responses (in session 1) used to establish the ground truth.

Data Analysis

Aware of the low number of participants, we tried to take every step to 1) maximize power by using all available data (from individual trials) with appropriate statistical tools and 2) ensure the robustness of results by cross-validating the findings across different measures and approaches.

Firstly, a manipulation check was carried out to ensure that our outcome variables were sensitive to the experimental manipulations, by testing the effect of the question phrasing (direct *vs.* indirect) and condition (polygraph *vs.* interrogation) on the outcome variables. This analysis was performed using mixed models with the participants and questions both entered as random factors. Marginal contrasts analysis (denoted by Δ) was also performed to clarify the differences between conditions. To allow for a better quantification of the uncertainty associated with the effects, as well as to increase the robustness to outliers and artefactual findings, all statistics were undertaken under the Bayesian framework (Makowski et al., 2019), using informative priors centred around 0 ($t_{Confidence}(1, 0, 1)$, $t_{RT}(1, 0, 3)$, $t_{Heartrate}(1, 0, 8)$).

To maximize the signal-to-noise ratio, we performed a feature reduction on our two groups of predictor variables (namely, ToM and interoception) using factor analysis over PCA, as the goal was to extract meaningful and consistent factors, rather than merely maximizing the variance explained. Then, we modelled the relationship between these inter-individual composite scores (note that the analysis for all individual variables is nonetheless included in the analysis report) and the 3 outcome variables in interaction with the condition (polygraph *vs.* interrogation). Finally, we investigated the relationship between the deception scale traits, and the ToM and interoception scores using Bayesian correlations. As all the analyses and data have been made available, we will in the manuscript focus on significant, i.e., - in this context - statistically reliable and in our opinion theoretically relevant.

The data analysis was carried out using *R 4.2* (R Core Team, 2022), *brms* (Bürkner, 2017), and the *easystats* collection of packages (Lüdecke et al., 2019, 2021; Makowski et al., 2019, 2020), and the physiological signal processing was done using the default routines available in *NeuroKit2* (Makowski, Pham, Lau, Brammer, et al., 2021). Note that EDA was not further analyzed as most participants did not yield any skin conductance responses - which we believe was partly caused by the low temperature (with dry air-con air) of the experimental room.

The analysis was not pre-registered (stemming out from an undergraduate's final year project), but the full reproducible analysis script, statistical results report, and data, are available at [**masked for blinding**]

Results

Manipulation Check

Compared to truth, lies were rated with less confidence ($\Delta = -1.35$, 95% CI $[-1.46, -1.23]$, $pd =$ 100%), but no significant difference between the conditions was found. On the other hand, the RT did not differ between truth and lies, but was significantly slower in the polygraph condition for both conditions ($\Delta = 0.25$, 95% CI [0.62, 0.41], pd = 100%). The heart rate was significantly more elevated during lies as compared to truth ($\Delta = 1.16$, 95% CI [0.57, 1.73], pd = 100%), and during interrogation as compared to the polygraph condition ($\Delta = 4.84$, 95% CI [4.23, 5.44], pd = 100%).

The indirect phrasing of the question only had a significant effect on RT ($\beta = 0.36$, 95% CI [0.21, 0.51], pd = 100%), leading to slower answers, regardless of whether they were lies or truths. Given this absence of interaction with the type of answers in any modality, this factor was not included in subsequent analysis.

Feature Reduction

The 3 Yoni-task dimensions and the 2 BES traits were combined into a unique factor, labelled *ToM* (explaining 35.76% of variance). It was loaded by the cognitive (.89), affective (.77), physical (.45) Yoni dimensions, and the affective (.41) and cognitive (.17) facets of the BES.

The 8 MAIA dimensions and the 3 HCT components were reduced to 4 factors (explaining 65.17% of variance). The first factor, labelled *Interoception - Meta* (23.59%), was loaded primarily by Attention Regulation (.97), Self-regulation (.63), Emotional awareness (.60), and Noticing (.49) dimensions of the MAIA and the HCT confidence score (.40). The second factor, labelled *Interoception - Listening* (18.54%), was primarily loaded by the Body Listening (.92) and Trusting (.53) MAIA dimensions, and the Awareness (-.60) and Confidence (.46) HCT scores. The third factor, labelled *Interoception - Focus* (12.07%), was primarily loaded by MAIA Not-Distracting (.87), Emotional Awareness (-.40) and HCT Accuracy (.33). The fourth factor, labelled *Interoception - Regulation* (10.97%), was primarily loaded by MAIA not-worrying (.71), HCT Accuracy (.61) and MAIA Trusting (.40).

Theory of Mind

The composite *ToM* score was significantly associated with less confident ($\beta = -0.19$, 95% CI [-0.36, -0.02], pd = 98.47%) and slower lies ($\beta = 0.42$, 95% CI [0.01, 0.83], pd = 97.67%), specifically in the polygraph condition. No significant effect was found with regards to heart rate, and no correlation was observed with dispositional lying traits (**Figure 1**).

Interoception

The *Meta* interoception score was significantly associated with an increased confidence in lies, specifically in the polygraph condition ($\beta = 0.20,\ 95\%\ CI\ [0.03, 0.35],\ pd = 98.98\%$). It was also associated with faster answers for both lies ($\beta = -0.54,\ 95\%\ CI\ [-0.93, -0.15],\ pd = 99.67\%$) and truths ($\beta = -0.29,\ 95\%\ CI\ [-0.63, 0.03],\ pd = 96.10\%$), specifically in the polygraph condition. No significant association was found with heart rate.

The *Listening* interoception score was significantly associated with an increased confidence in lies, in both the polygraph ($\beta = 0.43,\ 95\%\ CI\ [0.27, 0.59],\ pd = 100\%$) and interrogation condition ($\beta = 0.16,\ 95\%\ CI\ [0.01, 0.32],\ pd = 98.04\%$). It was also associated with faster answers, particularly for lies ($\beta = -0.42,\ 95\%\ CI\ [-0.82, -0.03],\ pd = 98.19\%$) and truths ($\beta = -0.36,\ 95\%\ CI\ [-0.76, 0.03],\ pd = 96.49\%$) in the polygraph condition. No significant association was found with heart rate. This score also correlated with the dispositional lying *Contextuality* trait ($r = 0.50,\ 95\%\ CI\ [0.04, 0.64],\ BF_{10} = 3.48\%$).

The *Focus* interoception score was significantly associated with an increased confidence in

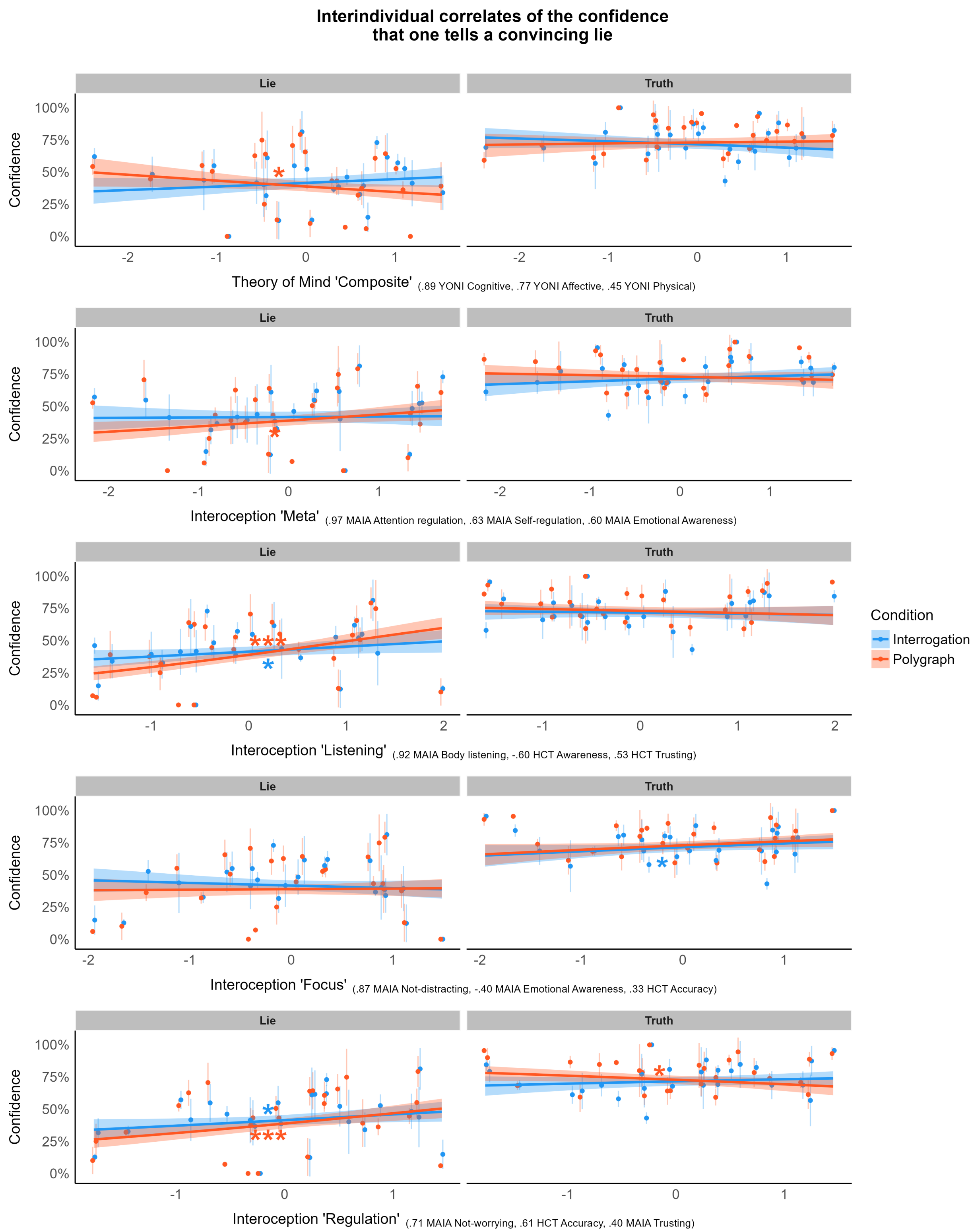

Figure 1. Interindividual corrrelates of lying confidence. The lines shows the relationship (with 95% CI uncertainty), assessed via Bayesian mixed models (**pd > 97%, **pd > 99%, ***pd > 99.9%), between the participants' interoceptive and ToM composite scores and the confidence ratings of their responses. Average lying confidence (+/- 1 SD) within the two experimental conditions is displayed as points for descriptive purposes as the models were ran on individual trials.

truthful responses in both the polygraph ($\beta = 0.17,\ 95\%\ CI\ [-0.01, 0.34],\ pd = 97.16\%$) and interrogation condition ($\beta = 0.15,\ 95\%\ CI\ [-0.02, 0.32],\ pd = 95.76\%$). No significant association was found with RT or heart rate, but the score correlated with the dispositional lying *Ability* trait ($r = 0.50,\ 95\%\ CI\ [0.22, 0.74],\ BF_{10} = 34.37\%$).

The *Regulation* interoception score was significantly associated with an increased confidence in lies in both the polygraph ($\beta = 0.32,\ 95\%\ CI\ [0.14, 0.51],\ pd = 99.99\%$) and the interrogation conditions ($\beta = 0.18,\ 95\%\ CI\ [0.00, 0.36],\ pd = 97.42\%$), and with a decreased confidence in truth only in the polygraph condition ($\beta = -0.1,\ 95\%\ CI\ [-0.36, 0.01],\ pd = 97.16\%$). No significant association was found with RT or heart rate (**Figure 2**).

Discussion

The present study examined the contribution of ToM and interoception on our ability to lie using a directed lie paradigm with two conditions ("Interrogation" and "Polygraph") designed to enhance each of the two mechanisms. Interestingly, we found that when participants were presented with (fake) physiological feedback (the polygraph condition), instead of a face of a person they had to lie to (the interrogation condition), their response time for both lies and truths increased, as did their heart rate. Although the condition did not impact the subjective confidence that participants had in their answers, it suggests that believing oneself to be submitted to a machine that is supposedly able to detect deception by interpreting physiological signals is a harder and/or more stressful condition than lying to a person. While research linking interoception and deception is limited, our results are in line with studies that show an association between interoceptive awareness and anxiety (Domschke et al., 2010; Garfinkel & Critchley, 2013; Yoris et al., 2015). Specifically, enhancing one's attention towards their internal bodily signals could have resulted in a hyper-vigilance towards physiological sensations that is perceived negatively. This is consistent with previous deception detection studies, in which participants' spontaneous lying behaviour only decreased when they were given feedback by a polygraph machine, but not when no feedback

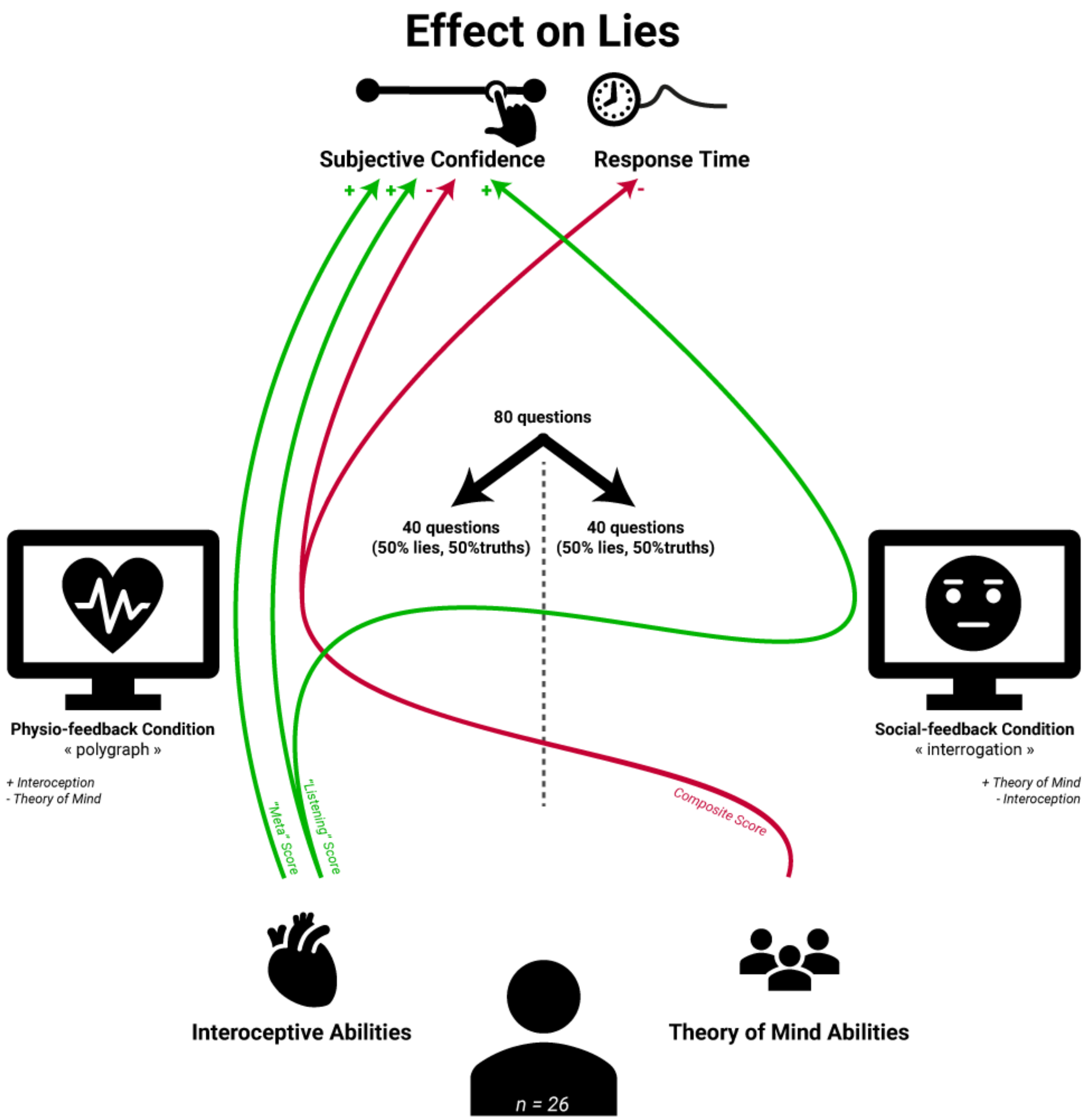

Figure 2. Summary of findings showing the positive (green) and negative (red) associations between interoception and theory of mind abilities and deception skills, depending on the experimental condition. It highlights that ToM was related to less confident and slower lies in the polygraph condition, and that specific interoceptive dimensions were related to more confident lies.

was given (Peleg et al., 2019). By extension, our study adds to the controversial discourse surrounding the use of physiological measures in past deception research, further questioning its validity as an indicator of deception (Oviatt et al., 2018; Rosky, 2013).

Our results suggest that higher ToM abilities were related to slower and less confident lies, but only in the polygraph condition. While previous bodies of work have reported mixed findings regarding the association between interoception and ToM (Chiou & Lee, 2013; Gendolla & Wicklund, 2009; Scaffidi Abbate et al., 2016; Wundrack & Specht, 2023), our results suggest the two are negatively linked. One possible interpretation of our findings is that people with stronger ToM abilities by default rely more on their social skills and altercentric inference when lying (i.e., they focus on - and try to read - the other person). When that mechanism is unavailable or unsuited (e.g., when there is no person to lie to - but a "machine" in our case), their corresponding lying ability decreases. However, in light of the current field of mixed findings relating interoception and ToM (Canino et al., 2022; Gao et al., 2019; Miller, 2015; Shah et al., 2017), future studies are necessary to investigate the interaction of these mechanisms in different social contexts.

We also showed that interoceptive abilities (as indicated by the composite interoception scores) are correlated with a higher confidence in one's lies in the polygraph condition, a condition in which the attention towards internal reactions is fostered. Indeed, this is in line with previous studies that found individuals with low interoception were more averse to risk when reputational stakes were high, telling fewer egoistical lies (Vabba et al., 2022). In fact, Vabba et al. (2022) further reports that people with high interoception abilities were less likely to differ in risk-taking tendencies, telling the same number of lies regardless of the social stakes. Consistent with our results, Mohr et al. (2023) further reports that individuals with high interoceptive accuracy were more likely to make egocentric decisions. However, in contrast to previous studies (Füstös et al., 2013; Owens et al., 2018; Pinna & Edwards, 2020; Pollatos et al., 2007), we did not find any significant relationship between individuals' interoception scores and their heart rate changes during their answers. This points toward a predominantly meta-cognitive effect without necessarily an actual bodily regulation (i.e., participants with good interoception feel that their lies are more convincing, but do not actively attenuate their bodily reactions).

Another possibility that should be tested in the future is that of a mediating role of executive functions, given their positive association with interoception (Molnar-Szakacs & Uddin, 2022). For instance, neuroscientific findings investigating the correlates of interoception have underlined the potential role of the anterior cingulate cortex (ACC) and anterior insula (AI) (Craig, 2009; Critchley et al., 2004; Khalsa et al., 2009; Wang et al., 2019), both of which are often thought to be activated during deception (Abe, 2011; Baumgartner et al., 2013; Sip et al., 2008), and have been implicated in cognitive processes associated with deception (such as cognitive control, Molnar-Szakacs & Uddin, 2022; or conflict detection, Kerns et al., 2004). It is thus possible that the positive relationship between interoceptive abilities and deception is at least partially mediated by cognitive control abilities.

Although yielding promising results, the sample size of this exploratory study is a source of concern. Although we tried to mitigate it by 1) extracting more robust variables (by combining multiple ones by means of feature reduction) and 2) using a suited analysis approach (Bayesian statistics with informative priors), future replication studies with larger samples are warranted to confirm this first investigation. Nonetheless, we believe our results to be credible as we find consistent patterns across various facets and measures (for instance, all interoceptive dimensions, although distinct, share a similar trend) in line with theoretical expectations. The statistical power could also explain the overall lack of results found in relation to heart rate, which has a higher signal-to-noise ratio as compared to subjective reports (such as confidence scales). Additionally, one has to note that the participants did not have strong incentive for lying (there was no risk of losing the “reward” - i.e., student credits), which might have further decreased the potential effect sizes.

Another aspect to note is the strong reliance on self-reported measures as outcome variables of lie ability (in particular, the measure of answer confidence, but also the auto-questionnaires). This might conflate meta-cognitive abilities as well as dishonest answers. Although we tried to include more objective measures, such as RT (although it too was tied to the participants' conscious decision to press a key) and heart rate, future studies should attempt at measuring objectively the answer (lie or truth) quality, for instance by means of external examiners. Note that this is not a limitation *per se*, as it answers a slightly different question - what are the correlates of *objective* lying skills - than deception self-confidence.

Additionally to the limitations pertaining to the measure of lying ability, some also concern the measure of the predictor constructs, namely ToM and interoception. While we tried to include a behavioral task as well as a subjective questionnaire for each, it has to be underlined that they are notoriously difficult concepts to measure. In particular, objective interoceptive accuracy was assessed using the Heartbeat Counting Task (HCT). While the HCT used to be considered as a gold standard and remains one of the most commonly used measures (Desmedt et al., 2022), concerns regarding its validity has been highlighted in several studies (Brener & Ring, 2016; Desmedt et al., 2018, 2022; Legrand et al., 2022). Given increasing research efforts invested in the development of novel interoception tasks (Legrand et al., 2022; Plans et al., 2021; Ponzo et al., 2021), future works should further examine the relationship between interoception and lying ability using measures with better psychometric properties.

Moreover, although the cognitive and affective components of ToM and empathy share overlaps in the current literature, and there is no consensus regarding how the two concepts should be delineated, recent evidence nonetheless suggests ToM and empathy are necessarily distinct constructs with separable underlying mechanisms (Kanske et al., 2015). As such, future studies are warranted to further investigate the associations between ToM and lying ability using validated instruments sensitive to measuring ToM (such as the Theory of Mind

Inventory, Hutchins et al., 2021). Furthermore, our application of feature reduction as a noise-elimination measure could have over-simplified the data. A more complex pattern of relationships, with different contributions of various subdimensions of ToM and interoception, could emerge provided a sufficient statistical power and valid measures.

In conclusion, this study is a first step towards assessing the contribution of ToM and interoception abilities in deception, particularly in one's ability to lie convincingly. To this end, we introduced a new paradigm to delineate the contribution of these mechanisms while remaining relevant to applied fields of lie detection and criminology (in which the experimental conditions find echoing practices). Notably, our results provide some evidence that interoception could be an important - and overlooked - process involved in deception. Furthermore, our findings extend and offer an alternate perspective to the debatable use of polygraphs, suggesting that its utility for lie detection is not only questionable, but could potentially selectively modulate deceptive skills depending on the cognitive and interoceptive profile of the participant.

Data Availability

The material (stimuli generation code, experiment code, raw data, analysis script with complementary figures and analyses, etc.) for this research is available at [**masked for blinding**].

Conflict of Interest Statement

The authors declare no conflict of interest.

Acknowledgements

[**masked for blinding**]

References

Abe, N. (2011). How the brain shapes deception: An integrated review of the literature. *The Neuroscientist*, *17*(5), 560–574.

Baron-Cohen, S. (1997). *Mindblindness: An essay on autism and theory of mind*. MIT press.

Barrett, L. F., & Simmons, W. K. (2015). Interoceptive predictions in the brain. *Nature Reviews Neuroscience*, *16*(7), 419–429.

Baumgartner, T., Gianotti, L. R., & Knoch, D. (2013). Who is honest and why: Baseline activation in anterior insula predicts inter-individual differences in deceptive behavior. *Biological Psychology*, *94*(1), 192–197.

Beaudoin, C., Leblanc, É., Gagner, C., & Beauchamp, M. H. (2020). Systematic review and inventory of theory of mind measures for young children. *Frontiers in Psychology*, *10*, 2905.

Bond Jr, C. F., & DePaulo, B. M. (2008). Individual differences in judging deception: Accuracy and bias. *Psychological Bulletin*, *134*(4), 477.

Bora, E., & Yener, G. G. (2017). Meta-analysis of social cognition in mild cognitive impairment. *Journal of Geriatric Psychiatry and Neurology*, *30*(4), 206–213.

Brener, J., & Ring, C. (2016). Towards a psychophysics of interoceptive processes: The measurement of heartbeat detection. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1708), 20160015.

Burgoon, J. K., & Buller, D. B. (1994). Interpersonal deception: III. Effects of deceit on perceived communication and nonverbal behavior dynamics. *Journal of Nonverbal Behavior*, *18*(2), 155–184.

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Canino, S., Raimo, S., Boccia, M., Di Vita, A., & Palermo, L. (2022). On the embodiment of social cognition skills: The inner and outer body processing differently contributes to the affective and cognitive theory of mind. *Brain Sciences*, *12*(11), 1423.

Chen, W. G., Schloesser, D., Arensdorf, A. M., Simmons, J. M., Cui, C., Valentino, R., Gnadt, J. W., Nielsen, L., Hillaire-Clarke, C. S., Spruance, V., et al. (2021). The emerging science of interoception: Sensing, integrating, interpreting, and regulating signals within the self. *Trends in Neurosciences*, *44*(1), 3–16.

Chiou, W.-B., & Lee, C.-C. (2013). Enactment of one-to-many communication may induce self-focused attention that leads to diminished perspective taking: The case of facebook. *Judgment and Decision Making*, *8*(3), 372–380.

Connell, L., Lynott, D., & Banks, B. (2018). Interoception: The forgotten modality in perceptual grounding of abstract and concrete concepts. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1752), 20170143.

Craig, A. D. (2009). How do you feel—now? The anterior insula and human awareness. *Nature Reviews Neuroscience*, *10*(1), 59–70.

Critchley, H. D., Wiens, S., Rotshtein, P., Öhman, A., & Dolan, R. J. (2004). Neural systems supporting interoceptive awareness. *Nature Neuroscience*, *7*(2), 189–195.

Damasio, A. R. (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *351*(1346), 1413–1420.

Desmedt, O., Luminet, O., & Corneille, O. (2018). The heartbeat counting task largely involves non-interoceptive processes: Evidence from both the original and an adapted counting task. *Biological Psychology*, *138*, 185–188.

Desmedt, O., Van Den Houte, M., Walentynowicz, M., Dekeyser, S., Luminet, O., & Corneille, O. (2022). How does heartbeat counting task performance relate to theoretically-relevant mental health outcomes? A meta-analysis. *Collabra: Psychology*, *8*(1), 33271.

Ditto, P. H., Pizarro, D. A., Epstein, E. B., Jacobson, J. A., & MacDonald, T. K. (2006). Visceral influences on risk-taking behavior. *Journal of Behavioral Decision Making*, *19*(2), 99–113.

Domschke, K., Stevens, S., Pfleiderer, B., & Gerlach, A. L. (2010). Interoceptive sensitivity in anxiety and anxiety disorders: An overview and integration of neurobiological findings. *Clinical Psychology Review*, *30*(1), 1–11.

Dunn, B. D., Galton, H. C., Morgan, R., Evans, D., Oliver, C., Meyer, M., Cusack, R., Lawrence, A. D., & Dalgleish, T. (2010). Listening to your heart: How interoception shapes emotion experience and intuitive decision making. *Psychological Science*, *21*(12), 1835–1844.

Evans, A. D., & Lee, K. (2011). Verbal deception from late childhood to middle adolescence and its relation to executive functioning skills. *Developmental Psychology*, *47*(4), 1108.

Furman, D. J., Waugh, C. E., Bhattacharjee, K., Thompson, R. J., & Gotlib, I. H. (2013). Interoceptive awareness, positive affect, and decision making in major depressive disorder. *Journal of Affective Disorders*, *151*(2), 780–785.

Füstös, J., Gramann, K., Herbert, B. M., & Pollatos, O. (2013). On the embodiment of emotion regulation: Interoceptive awareness facilitates reappraisal. *Social Cognitive and Affective Neuroscience*, *8*(8), 911–917.

Gallant, C. M., Lavis, L., & Mahy, C. E. (2020). Developing an understanding of others' emotional states: Relations among affective theory of mind and empathy measures in early childhood. *British Journal of Developmental Psychology*, *38*(2), 151–166.

Gao, Q., Ping, X., & Chen, W. (2019). Body influences on social cognition through interoception. *Frontiers in Psychology*, *10*, 2066.

Garfinkel, S. N., & Critchley, H. D. (2013). Interoception, emotion and brain: New insights link internal physiology to social behaviour. Commentary on: "Anterior insular cortex mediates bodily sensibility and social anxiety" by terasawa et al.(2012). *Social Cognitive and Affective Neuroscience*, *8*(3), 231–234.

Garfinkel, S. N., Seth, A. K., Barrett, A. B., Suzuki, K., & Critchley, H. D. (2015). Knowing your own heart: Distinguishing interoceptive accuracy from interoceptive awareness. *Biological Psychology*, *104*, 65–74.

Gendolla, G. H., & Wicklund, R. A. (2009). Self-focused attention, perspective-taking, and false consensus. *Social Psychology*, *40*(2), 66–72.

Gonzalez-Billandon, J., Aroyo, A. M., Tonelli, A., Pasquali, D., Sciutti, A., Gori, M., Sandini, G., & Rea, F. (2019). Can a robot catch you lying? A machine learning system to detect lies during interactions. *Frontiers in Robotics and AI*, *6*, 64.

Hutchins, T. L., Lewis, L., Prelock, P. A., & Brien, A. (2021). The development and preliminary psychometric evaluation of the theory of mind inventory: Self report—adult (ToMI: SR-adult). *Journal of Autism and Developmental Disorders*, *51*, 1839–1851.

Jolliffe, D., & Farrington, D. P. (2006). Development and validation of the basic empathy scale. *Journal of Adolescence*, *29*(4), 589–611.

Kanske, P., Böckler, A., Trautwein, F.-M., & Singer, T. (2015). Dissecting the social brain: Introducing the EmpaToM to reveal distinct neural networks and brain–behavior relations for empathy and theory of mind. *NeuroImage*, *122*, 6–19.

Kerns, J. G., Cohen, J. D., MacDonald III, A. W., Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science*, *303*(5660), 1023–1026.

Khalsa, S. S., Rudrauf, D., Feinstein, J. S., & Tranel, D. (2009). The pathways of interoceptive awareness. *Nature Neuroscience*, *12*(12), 1494–1496.

Kireev, M., Korotkov, A., Medvedeva, N., & Medvedev, S. (2013). Possible role of an error detection mechanism in brain processing of deception: PET-fMRI study. *International Journal of Psychophysiology*, *90*(3), 291–299.

Lee, J. Y. S., & Imuta, K. (2021). Lying and theory of mind: A meta-analysis. *Child Development*, *92*(2), 536–553.

Legrand, N., Nikolova, N., Correa, C., Brændholt, M., Stuckert, A., Kildahl, N., Vejlø, M., Fardo, F., & Allen, M. (2022). The heart rate discrimination task: A psychophysical method to estimate the accuracy and precision of interoceptive beliefs. *Biological Psychology*, *168*, 108239.

Lenggenhager, B., Azevedo, R. T., Mancini, A., & Aglioti, S. M. (2013). Listening to your heart and feeling yourself: Effects of exposure to interoceptive signals during the ultimatum game. *Experimental Brain Research*, *230*(2), 233–241.

Levine, E. E., & Lupoli, M. J. (2022). Prosocial lies: Causes and consequences. *Current Opinion in Psychology*, *43*, 335–340.

Levine, E. E., & Schweitzer, M. E. (2014). Are liars ethical? On the tension between benevolence and honesty. *Journal of Experimental Social Psychology*, *53*, 107–117.

Levine, E. E., & Schweitzer, M. E. (2015). Prosocial lies: When deception breeds trust. *Organizational Behavior and Human Decision Processes*, *126*, 88–106.

Levine, T. R., Serota, K. B., Shulman, H., Clare, D. D., Park, H. S., Shaw, A. S., Shim, J. C., & Lee, J. H. (2011). Sender demeanor: Individual differences in sender believability have a powerful impact on deception detection judgments. *Human Communication Research*, *37*(3), 377–403.

Lüdecke, D., Ben-Shachar, M., Patil, I., Waggoner, P., & Makowski, D. (2021). performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software*, *6*(60), 3139. https://doi.org/10.21105/joss.03139

Lüdecke, D., Waggoner, P., & Makowski, D. (2019). Insight: A unified interface to access information from model objects in R. *Journal of Open Source Software*, *4*(38), 1412. https://doi.org/10.21105/joss.01412

Makowski, D., Ben-Shachar, M. S., & Lüdecke, D. (2019). bayestestR: Describing effects and their uncertainty, existence and significance within the bayesian framework. *Journal of Open Source Software*, *4*(40), 1541.

Makowski, D., Ben-Shachar, M., Patil, I., & Lüdecke, D. (2020). Methods and algorithms for correlation analysis in R. *Journal of Open Source Software*, *5*(51), 2306. https://doi.org/10.21105/joss.02306

Makowski, D., Pham, T., Lau, Z. J., Brammer, J. C., Lespinasse, F., Pham, H., Schölzel, C., & Chen, S. (2021). NeuroKit2: A python toolbox for neurophysiological signal processing. *Behavior Research Methods*, *53*(4), 1689–1696.

Makowski, D., Pham, T., Lau, Z. J., Raine, A., & Chen, S. (2021). The structure of deception: Validation of the lying profile questionnaire. *Current Psychology*, 1–16.

Masip, J. (2017). Deception detection: State of the art and future prospects. *Psicothema*, *29*(2), 149–159.

Mehling, W. E., Price, C., Daubenmier, J. J., Acree, M., Bartmess, E., & Stewart, A. (2012). The multidimensional assessment of interoceptive awareness (MAIA). *PloS One*, *7*(11), e48230.

Miller, J. E. (2015). *The connections between self-monitoring and theory of mind*.

Mohr, M. von, Finotti, G., Esposito, G., Bahrami, B., & Tsakiris, M. (2023). Social interoception: Perceiving events during cardiac afferent activity makes people more suggestible to other people's influence. *Cognition*, *238*, 105502.

Molnar-Szakacs, I., & Uddin, L. Q. (2022). Anterior insula as a gatekeeper of executive control. *Neuroscience & Biobehavioral Reviews*, 104736.

Murphy, J., Catmur, C., & Bird, G. (2019). Classifying individual differences in interoception: Implications for the measurement of interoceptive awareness. *Psychonomic Bulletin & Review*, *26*(5), 1467–1471.

Oviatt, S., Schuller, B., Cohen, P. R., Sonntag, D., Potamianos, G., & Krüger, A. (2018). *The handbook of multimodal-multisensor interfaces: Signal processing,* *architectures, and detection of emotion and cognition-volume 2.* Association for Computing Machinery; Morgan & Claypool.

Owens, A. P., Friston, K. J., Low, D. A., Mathias, C. J., & Critchley, H. D. (2018). Investigating the relationship between cardiac interoception and autonomic cardiac control using a predictive coding framework. *Autonomic Neuroscience*, *210*, 65–71.

Paulhus, D. L., & Williams, K. M. (2002). The dark triad of personality: Narcissism, machiavellianism, and psychopathy. *Journal of Research in Personality*, *36*(6), 556–563.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203.

Peleg, D., Ayal, S., Ariely, D., & Hochman, G. (2019). The lie deflator-the effect of polygraph test feedback on subsequent (dis) honesty. *Judgment & Decision Making*, *16*(6).

Petzschner, F. H., Garfinkel, S. N., Paulus, M. P., Koch, C., & Khalsa, S. S. (2021). Computational models of interoception and body regulation. *Trends in Neurosciences*, *44*(1), 63–76.

Pinna, T., & Edwards, D. J. (2020). A systematic review of associations between interoception, vagal tone, and emotional regulation: Potential applications for mental health, wellbeing, psychological flexibility, and chronic conditions. *Frontiers in Psychology*, *11*, 1792.

Plans, D., Ponzo, S., Morelli, D., Cairo, M., Ring, C., Keating, C. T., Cunningham, A., Catmur, C., Murphy, J., & Bird, G. (2021). Measuring interoception: The phase adjustment task. *Biological Psychology*, *165*, 108171.

Pollatos, O., Herbert, B. M., Matthias, E., & Schandry, R. (2007). Heart rate response after emotional picture presentation is modulated by interoceptive awareness. *In-*

ternational Journal of Psychophysiology, *63*(1), 117–124.

Ponzo, S., Morelli, D., Suksasilp, C., Cairo, M., & Plans, D. (2021). Measuring interoception: The CARdiac elevation detection task. *Frontiers in Psychology*, *12*.

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rassin, E., Sergiou, C., Linden, D. van der, & Dongen, J. van. (2023). Psychopathy as a predisposition to lie hedonistically. *Psychology, Crime & Law*, 1–8.

Riggio, R. E., Tucker, J., & Widaman, K. F. (1987). Verbal and nonverbal cues as mediators of deception ability. *Journal of Nonverbal Behavior*, *11*, 126–145.

Roheger, M., Brenning, J., Riemann, S., Martin, A. K., Flöel, A., & Meinzer, M. (2022). Progression of socio-cognitive impairment from healthy aging to alzheimer's dementia: A systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews*, 104796.

Rosky, J. W. (2013). The (f) utility of post-conviction polygraph testing. *Sexual Abuse*, *25*(3), 259–281.

Rubin, D. C., Schrauf, R. W., & Greenberg, D. L. (2003). Belief and recollection of autobiographical memories. *Memory & Cognition*, *31*, 887–901.

Scaffidi Abbate, C., Boca, S., & Gendolla, G. H. (2016). Self-awareness, perspective-taking, and egocentrism. *Self and Identity*, *15*(4), 371–380.

Schandry, R. (1981). Heart beat perception and emotional experience. *Psychophysiology*, *18*(4), 483–488.

Sebastian, C. L., Fontaine, N. M., Bird, G., Blakemore, S.-J., De Brito, S. A., McCrory, E. J., & Viding, E. (2012). Neural processing associated with cognitive and affective theory of mind in adolescents and adults. *Social Cognitive and Affective Neuroscience*, *7*(1), 53–63.

Seth, A. K., Suzuki, K., & Critchley, H. D. (2012). An interoceptive predictive coding model of conscious presence. *Frontiers in Psychology*, *2*, 395.

Shah, P., Catmur, C., & Bird, G. (2017). From heart to mind: Linking interoception, emotion, and theory of mind. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *93*, 220.

Shamay-Tsoory, S. G., & Aharon-Peretz, J. (2007). Dissociable prefrontal networks for cognitive and affective theory of mind: A lesion study. *Neuropsychologia*, *45*(13), 3054–3067.

Shamay-Tsoory, S. G., Harari, H., Aharon-Peretz, J., & Levkovitz, Y. (2010). The role of the orbitofrontal cortex in affective theory of mind deficits in criminal offenders with psychopathic tendencies. *Cortex*, *46*(5), 668–677.

Sip, K. E., Roepstorff, A., McGregor, W., & Frith, C. D. (2008). Detecting deception: The scope and limits. *Trends in Cognitive Sciences*, *12*(2), 48–53.

Sip, K. E., Skewes, J. C., Marchant, J. L., McGregor, W. B., Roepstorff, A., & Frith, C. D. (2012). What if i get busted? Deception, choice, and decision-making in social interaction. *Frontiers in Neuroscience*, *6*, 58.

Sternglanz, R. W., Morris, W. L., Morrow, M., & Braverman, J. (2019). A review of meta-analyses about deception detection. *The Palgrave Handbook of Deceptive Communication*, 303–326.

Sugawara, A., Terasawa, Y., Katsunuma, R., & Sekiguchi, A. (2020). Effects of interoceptive training on decision making, anxiety, and somatic symptoms. *BioPsychoSocial Medicine*, *14*, 1–8.

Takuma, K., Hori, S., Sasaki, J., Shinozawa, Y., Yoshikawa, T., Handa, S., Horikawa, M., & Aikawa, N. (1995). An alternative limb lead system for electrocardiographs in emergency patients. *The American Journal of Emergency Medicine*, *13*(5), 514–517. https://doi.org/10.1016/0735-6757(95)90160-4

Talwar, V., Crossman, A., & Wyman, J. (2017). The role of executive functioning and theory of mind in children's lies for another and for themselves. *Early Childhood*

Research Quarterly, *41*, 126–135.

Talwar, V., Gordon, H. M., & Lee, K. (2007). Lying in the elementary school years: Verbal deception and its relation to second-order belief understanding. *Developmental Psychology*, *43*(3), 804.

Vabba, A., Porciello, G., Panasiti, M. S., & Aglioti, S. M. (2022). Interoceptive influences on the production of self-serving lies in reputation risk conditions. *International Journal of Psychophysiology*, *177*, 34–42.

Verigin, B. L., Meijer, E. H., Bogaard, G., & Vrij, A. (2019). Lie prevalence, lie characteristics and strategies of self-reported good liars. *PloS One*, *14*(12), e0225566.

Viji, D., Gupta, N., & Parekh, K. H. (2022). History of deception detection techniques. *Proceedings of International Conference on Deep Learning, Computing and Intelligence*, 373–387.

Walczyk, J. J., & Cockrell, N. F. (2022). To err is human but not deceptive. *Memory & Cognition*, *50*(1), 232–244.

Walczyk, J. J., Harris, L. L., Duck, T. K., & Mulay, D. (2014). A social-cognitive framework for understanding serious lies: Activation-decision-construction-action theory. *New Ideas in Psychology*, *34*, 22–36.

Walczyk, J. J., Mahoney, K. T., Doverspike, D., & Griffith-Ross, D. A. (2009). Cognitive lie detection: Response time and consistency of answers as cues to deception. *Journal of Business and Psychology*, *24*, 33–49.

Wang, X., Wu, Q., Egan, L., Gu, X., Liu, P., Gu, H., Yang, Y., Luo, J., Wu, Y., Gao, Z., et al. (2019). Anterior insular cortex plays a critical role in interoceptive attention. *Elife*, *8*, e42265.

Weiss, S., Sack, M., Henningsen, P., & Pollatos, O. (2014). On the interaction of self-regulation, interoception and pain perception. *Psychopathology*, *47*(6), 377–382.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, *72*(3), 655–684.

Williams, E. F., Pizarro, D., Ariely, D., & Weinberg, J. D. (2016). The valjean effect: Visceral states and cheating. *Emotion*, *16*(6), 897.

Wundrack, R., & Specht, J. (2023). Mindful self-focus–an interaction affecting theory of mind? *Plos One*, *18*(2), e0279544.

Yoris, A., Esteves, S., Couto, B., Melloni, M., Kichic, R., Cetkovich, M., Favaloro, R., Moser, J., Manes, F., Ibanez, A., et al. (2015). The roles of interoceptive sensitivity and metacognitive interoception in panic. *Behavioral and Brain Functions*, *11*(1), 1–6.

Zvi, L., & Elaad, E. (2018). Correlates of narcissism, self-reported lies, and self-assessed abilities to tell and detect lies, tell truths, and believe others. *Journal of Investigative Psychology and Offender Profiling*, *15*(3), 271–286.