**The Heart Lies: The Role of Interoception and Theory of Mind in Deception**

Dominique Makowski^ 1, \*^, Zen J. Lau^ 1^, Tam Pham^ 1^, & S.H. Annabel Chen^ 1, 2, 3, 4^

1 School of Social Sciences, Nanyang Technological University, Singapore

2 Centre for Research and Development in Learning, Nanyang Technological University, Singapore

3 Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore

4 National Institute of Education, Nanyang Technological University, Singapore

Author note

Correspondence concerning this article should be addressed to S.H. Annabel Chen, HSS 04-19, 48 Nanyang Avenue, Singapore. E-mail: [annabelchen@ntu.edu.sg](mailto:annabelchen@ntu.edu.sg)

**The Heart Lies: The Role of Interoception and Theory of Mind in Deception**

# Introduction

The fabrication and delivery of false information is a sophisticated, and cognitively taxing activity (Abe et al., 2007; Leal et al., 2008 ), resulting in physiological arousal (Eskritt et al., 2021) which may require effortful self-regulation to conceal(Krokoszinski & Hosser, 2016).

Much of the current research investigating the process and mechanisms involved in lying has considered capabilities such as Theory of Mind (ToM), the ability to understand and make inferences about other people’s mental states and predict their behaviours (El Haj et al., 2017; Lee & Imuta, 2021). ToM is a critical component of deception, since the liar in question must be able to use their inferences to predict whether the individual accepted as true the false information presented to them (Stewart et al., 2019). Children begin using deception during the same period of development in which executive function and ToM develops (Sai et al., 2021), and children trained in ToM saw an increase in lying behaviour (Ding et al., 2015). However, there is a growing link between other cognitive and physiological mechanisms that may play a role in both deception detection, and fabricating successful lies

Interoceptive awareness (IA) is the ability to monitor, interpret, integrate, and regulate the sensations and state of one’s internal organs, and is closely tied to emotional and cognitive processes (Chen et al., 2021; Garfinkel et al., 2015). Greater is IA associated with greater self-regulation (Weiss et al., 2014), especially mindfulness practitioners who tend to have greater interoceptive awareness and greater ability to regulate stress and associated physiological indicators (Ardi et al., 2021; McMahon, 2021).

The growing literature examining the relationship between interoception and deception is rapidly growing area of inquiry. Some research has examined how higher interoceptive accuracy, measured through heart beat detection, has been associated with vasoconstriction when faced with a liar, indicating that IA may aid in discriminating high stakes emotional lies and truths (Gunderson et al., 2021). It is clear that interoception may indicate on an implicit level that one is being lied to (Ten Brinke et al., 2019; Veer et al., 2014), however, there is limited research investigating how interoception may work the other way around, in favour of a liar in regulating their own physiological state that may betray their deception.

Attending to interoceptive cues has been associated with greater willingness to engage in cheating (Williams et al., 2016), and risk-taking behaviours (Ditto et al., 2006; Lenggenhager et al., 2013). The findings by Vabba et al. (2022) indicate that high cardiac interoceptive participants were more likely to be dishonest, and to tell higher egoistic lies. There is limited research examining whether individuals with high interoceptive accuracy may be more confident in managing physiological tells associated with emotions like guilt or anxiety that the lie may be discovered through suppression or reappraisal (Füstös et al., 2013; Megı́as-Robles et al., 2019).

The act of lying is often accompanied by physiological responses such as increased heart rate, increased respiration and sweating. Since fluctuations in key physiological indicators as a means to determine whether the individual is lying is a key principle behind the polygraph test and other lie detection techniques, there is a need for more research examining the relationship between IA, theory of mind, and deception. This research attempts to isolate the effects of theory of mind and interoception, and examine how they interact with deception. Of particular interest is how liars might utilize their interoceptive ability to self regulate in towards a more successful deception, and how this ability may compare or interact with other mechanisms at play such as theory of mind. We hypothesize that participants with greater theory of mind ability will show higher lie confidence, shorter reaction time, and lower physiological arousal in the interrogation condition, while participants in the polygraph condition will demonstrate lower lie confidence, lower reaction time and higher physiological arousal during the polygraph prediction.

# Methods

The plan for this study was preregistered (OSF?). The raw data and analysis script can be found at (GITHUB?).

## Participants

*26* participants (65.4% Females, M*age* = 20.9, *SD* = 2.0) were

recruited VIA HOW

for HUAT

-exclusion criteria

This research was approved by the Institutional Review Board (Reference Number: IRB-) of Nanyang Technological University (NTU).

volunteered for [duration of study] for [compensation provided].

## Measures

### Theory of Mind and Empathy.

Since Theory of Mind and empathy are closely related, *The Basic Empathy Scale (BES)* (Jolliffe & Farrington, 2006) was used as a subjective measure for theory of mind. *BES* uses a 40-item Likert scale to rate statements such as, “I can usually work out when people are cheerful” from 1-5. We also implemented the *Yoni Task,* which assesses the ability to judge mental states based on verbal cues, eye gaze and facial expression (Shamay-Tsoory & Aharon-Peretz, 2007).

### Interoceptive Ability.

To measure interoception, participants completed a *Heartbeat Counting Task (HCT)* (Schandry, 1981), where participants count their heart beats without taking their pulse, as well as provide a confidence rating for their estimate. At the same time their true heart rate is recorded and the scores are compared to determine the individual’s accuracy. They also completed the *MAIA questionnaire* (Mehling et al., 2012), where they rated 32 statements such as “I can use my breath to reduce tension” on a 1-5 Likert scale.

Additionally, we implemented the Heartbeat Tracking Task (HTT), in which participants - tap one

### Deception.

We used the *LIE Scale*, which is a 44-item questionnaire examining deception as a dispositional trait, and measures the participant propensity towards lie frequency, ability, contextuality and negativity (Makowski et al., 2021).

#### Deception Task.

Participants were randomly allocated(??) between two conditions where they underwent 40 trials of a directed lying task. During the task, participants would be instructed to lie for 20 of the trials to questions such as, “What secondary school did you go to?”. Prior to beginning the experiment, researchers collected the answers via questionnaire. In each condition, participants received false feedback through different mediums according to the condition that they were assigned, intended to examine the relationship between theory of mind or interoception and the dependent variables (see fig. X?). Participants were randomized between two conditions which we titled the Interrogation Condition and the Polygraph Condition.

In the Interrogation condition, participants were instructed that following their lie, a live examiner over video feed would appear briefly to judge whether they were truthful or not. In actuality, the video of the examiner shown was pre-recorded and the same for all participants. We hypothesized that participants would will rely on theory of mind abilities to predict whether their lie to the examiner was successfully accepted as true.

The Polygraph Condition followed the same routine, however instead of a receiving feedback from an interrogator, following the lie, a video of physiological signals was shown. We hypothesized when made acutely aware of such signals through false feedback, that participants would rely on interoceptive abilities to be more mindful or intentional about modulating the signals that may give them away. The physiological signals in this condition are also false feedback, and pre-recorded and identical for all participants.

During each trial, the time it took for participants to answer is recorded, and their heart rates are measured using an electrocardiogram. Following the trial, participants gave a confidence rating for how successful they believe their lie or truth was.

During the interrogation condition, participants were exposed to video feedback of an examiner following each question. They were informed that the examiner was live and monitoring their physiological signals remotely to determine the honesty of their answers, however this was false, and the same sequence of videos was show to each participant.

### Demographic Information.

The demographic information we collected was age, ethnicity, nationality, dominant hand (for application of EDA electrodes), education level (determined by highest qualification received, or in the midst of achieving), religion and duration lived in Singapore.

## Procedure

Each subject was briefed on the experimental procedure and study aims and signed an informed consent document prior to the study commencing.

questionnaire to obtain truthful answers to directed lying task,

when were the measures like yoni etc taken?

sound attenuated room

Physio recording devices (equipment by huat and who)

Where/how did they provide their answer during directed lying task

Deception task

During the experiment, all participants were connected to an electrocardiogram to monitor their heart would be asked a question, [HOW WAS ANSWER GIVEN?] and then connected for 10 seconds to the condition-unique stimulus. Following the exposure to the stimulus, they would be asked to rate their confidence on a sliding scale before repeating the trial 40 times.

Debriefing was provided following the study.

## Data Analysis

The manipulation checks consist of testing whether there is an effect of question phrasing (direct vs. indirect), and condition (polygraph *vs.* interrogation), on our 3 outcome variables: the participants’ ratings of confidence that their answers (lies *vs.* truths) were convincing, the response time (RT), and the heart rate change associated with the response. This analysis was performed using mixed models with the participants and questions both entered as random factors. Marginal contrasts analysis was also performed to clarify the differences between conditions. To allow for a better quantification of the uncertainty associated with the effects, as well increase the robustness to outliers and artifactual findings, all statistics were undertaken under the Bayesian framework [**ref makowski2019existence**], using informative priors centred around 0 (, , ).

To further counterbalance the low number of participants and maximize the signal-to-noise ratio, we performed a feature reduction on our two groups of predictor variables (namely, theory of mind and interoception). We used factor analysis (using the method agreement procedure to estimate the optimal number of dimensions, **REF PARAMETERS PKG**), rather 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 (the analysis for all the variables is included in Supplementary Materials) and the 3 outcome variables in interaction with the condition (polygraph *vs.* interrogation). Finally, we also investigated the relationship between the deception scale traits, and the theory of mind and interoception scores using Bayesian correlations.

The data analysis was carried out using R, the *brms* package [*REF*] and the *easystats* ecosystem **[REFS]**. As the full reproducible analysis script and statistical results are available at **[INSERT LINK]**, we will only focus on significant results in the manuscript.

# Results

## Manipulation Check

Compared to truth, lies were rated with less confidence (). Moreover, lies told in the polygraph condition were rated with less confidence (). The RT did not differ between truth and lies, but was significantly slower in the polygraph condition () for both conditions. The heart rate was significantly more elevated during lies as compared to truth (), and during interrogation as compared to the polygraph condition ().

The indirect phrasing of the question only had a significant effect on RT (), leading to slower answers, regardless of whether they were lies or truths.

## Feature Reduction

The 3 YONI-task dimensions and the 2 BES traits were combined into a unique factor, labelled *Theory of Mind* (explaining 35.76% of variance). It was loaded by the cognitive (.89), affective (.77), physical (.45) YONI dimension, and the affective (.41) and cognitive (.17) traits 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). However, due to the interpretation difficulty, we will refer to these dimensions with their main correlates in **Figure 1**.

## Theory of Mind

The composite *ToM* score was significantly associated with less confident () and slower lies (), specifically in the polygraph condition. No significant effect was found heart rate, and no correlation with the deception scale traits.

## Interoception

The *Meta* interoception score was significantly associated with an increased confidence in lies, specifically in the polygraph condition (). It was also associated with faster answers for both lies () and truths (), 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 the polygraph () and interrogation condition (). It was also associated with faster answers, particilarly for lies () and truths () in the polygraph condition. No significant association was found with heart rate. This score also correlated with the deception *Contextuality* trait ().

The *Focus* interoception score was significantly associated with an increased confidence in truthful responses in the polygraph () and interrogation condition (). No significant association was found with RT or heart rate, but the score correlated with the deception *Ability* trait ().

The *Regulation* interoception score was significantly associated with an increased confidence in lies in the polygraph () and the interrogation conditions (), and with a decreased confidence in truth only in the polygraph condition (). No significant association was found with RT or heart rate.

# Discussion

This study aimed to examine the relationship between Theory of Mind and interoception on lie confidence, reaction time and physiological arousal during deception. This was achieved through a directed lying task in which participants were randomized between two two conditions of false feedback where they were intended to rely on either interoceptive or ToM abilities. Across both conditions, the scores show a larger effect on confidence in the polygraph condition, which was both expected and unexpected. This may be due to the polygraph condition may be harder due to the less intuitive feedback cues. Those better at interoception or perceiving their internal signals would hence be more confident in their lies, as we expected. On the other hand, those with better theory of mind become less confident in the absence of another person from which to obtain feedback cues.

There was no significant relationship between reaction time and heart rate, however lie confidence is negatively correlated with reaction time, which is expected. Generally, participants were less confident in lies than truth, however is effect was greater in the polygraph condition, which is possibly due to less intuitive feedback cues. This effect carries over into reaction time, where participants were also slower to respond in the polygraph condition, which is consistent with less intuitive cues.

To our knowledge, this is the first study that has found a significant correlation between interoceptive awareness and lie confidence. Both interoceptive measures, HCT and MAIA, were strongly correlated. Participants with greater IA, also had greater lie confidence and lower truth confidence in the polygraph condition. The three sub-scales of MAIA which had a significant interaction were *Noticing, Body Listening* and *Attention Regulation*. Since participants were made aware of their cardiac activation through false feedback, we anticipated that those with greater accuracy would demonstrate greater lie confidence. Since research has shown deception is more likely when individuals judge a lower risk of confrontation or being caught (Sip et al., 2012, 2013), these results may offer support to the literature that implicates IA with immoral behaviour. Thus our research may be in support of the findings from Vabba et al. (2022), where IA was found to be a significant moderator in participant’s propensity toward deception and high egoist lies. Similar findings have also linked increased bodily awareness was linked to a greater likelihood of engaging in immoral behavior such as cheating (Williams et al., 2016), and risk-taking behaviours (Ditto et al., 2006; Lenggenhager et al., 2013). However, more research is required to examine this potential link.

A greater HCT score was associated with decreased lie confidence in the polygraph condition. Since HCT awareness is understood to be the correlation between HCT accuracy and HCT confidence, there are two potential explanations for this. First, a low level of awareness could mean that the participant accurately counted their heartbeats, but were not confidence; or vice versa, where they were inaccurate counting but expressed high confidence. The dual interpretations of this data may indicate its lack of reliability. More research is needed to understand the relationship between lie confidence, interoception and potential implications on lying behaviour.

HCT accuracy is negatively correlated with heart rate, but only in the interrogation condition. While it is expected that those with greater interoceptive ability would be better at perceiving and regulating physiological arousal (Megı́as-Robles et al., 2019), we expected to see this result in the polygraph condition which emphasizes interoceptive abilities. A potential explanation for this is that the presence of another person, whom they believe is observing them is understood to increase an individual’s interoceptive accuracy (Hazem et al., 2017; Isomura & Watanabe, 2020), which may in turn, cause the participant to attempt to regulate their cardiac response.

Greater ToM abilities negatively correlated with lie confidence, and saw an increase in truth confidence, but only in the polygraph condition. In the Yoni Task, this effect is driven by the cognitive and physical domains. BES saw the effect driven by both the cognitive and affective domain for lie confidence, meanwhile, truth confidence is only driven by the affective domain.

We anticipated that ToM measures would negatively correlate with physiological arousal, however while the Yoni Task had no association with heart rate, we found that BES scores were positively correlated with heart rate for both lies and truths, with a larger effect in the interrogation condition. Although initially unexpected, this finding is makes sense when interpreted in line with past research which finds a close positive correlation between empathy and increased cardiac response (Jauniaux et al., 2020; Lischke et al., 2018). The lack of correlation between heart rate and the Yoni task, indicates that the elevated cardiac response may be due exclusively to empathy, and attributable to ToM.

## Limitations and Future Research

Potential limitations worth noting in this research is that the interrogation and polygraph conditions may not bring out Theory of Mind and interoceptive abilities to the greatest degree. This must be considered as a potential limitation since the participants were still connected to the ECG in the interrogation condition, which may amplify their interoceptive awareness beyond what they may normally experience during a real-world act of deception. Furthermore, it is possible that there could be an interplay between Theory of Mind and interoception. More research is needed to investigate this relationship in the context of deception.

# Funding

# Author Contributions

# Conflict of Interest Statement

# Acknowledgements

# References

Abe, N., Suzuki, M., Mori, E., Itoh, M., & Fujii, T. (2007). Deceiving others: Distinct neural responses of the prefrontal cortex and amygdala in simple fabrication and deception with social interactions. *Journal of Cognitive Neuroscience*, *19*(2), 287–295.

Ardi, Z., Golland, Y., Shafir, R., Sheppes, G., & Levit-Binnun, N. (2021). The effects of mindfulness-based stress reduction on the association between autonomic interoceptive signals and emotion regulation selection. *Psychosomatic Medicine*, *83*(8), 852–862.

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.

Ding, X. P., Wellman, H. M., Wang, Y., Fu, G., & Lee, K. (2015). Theory-of-mind training causes honest young children to lie. *Psychological Science*, *26*(11), 1812–1821.

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.

El Haj, M., Antoine, P., & Nandrino, J. L. (2017). When deception influences memory: The implication of theory of mind. *Quarterly Journal of Experimental Psychology*, *70*(7), 1166–1173.

Eskritt, M., Fraser, B., & Bosacki, S. (2021). Did you just lie to me? Deception detection in face to face versus computer mediated communication. *The Journal of Social Psychology*, 1–14.

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.

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.

Gunderson, C., Brinke, L. ten, & Sokol-Hessner, P. (2021). *When the body knows: Interoceptive accuracy enhances physiological but not explicit differentiation between liars and truth-tellers*.

Hazem, N., George, N., Baltazar, M., & Conty, L. (2017). I know you can see me: Social attention influences bodily self-awareness. *Biological Psychology*, *124*, 21–29.

Isomura, T., & Watanabe, K. (2020). Direct gaze enhances interoceptive accuracy. *Cognition*, *195*, 104113.

Jauniaux, J., Tessier, M.-H., Regueiro, S., Chouchou, F., Fortin-Côté, A., & Jackson, P. L. (2020). Emotion regulation of others’ positive and negative emotions is related to distinct patterns of heart rate variability and situational empathy. *PloS One*, *15*(12), e0244427.

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

Krokoszinski, L., & Hosser, D. (2016). Emotion regulation during deception: An EEG study of imprisoned fraudsters. *Journal of Criminal Psychology*.

Leal, S., Vrij, A., Fisher, R. P., & Van Hooff, H. (2008). The time of the crime: Cognitively induced tonic arousal suppression when lying in a free recall context. *Acta Psychologica*, *129*(1), 1–7.

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

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.

Lischke, A., Pahnke, R., Mau-Moeller, A., Behrens, M., Grabe, H. J., Freyberger, H. J., Hamm, A. O., & Weippert, M. (2018). Inter-individual differences in heart rate variability are associated with inter-individual differences in empathy and alexithymia. *Frontiers in Psychology*, *9*, 229.

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.

McMahon, S. (2021). *Mindful success: A correlational study on mindfulness, interoceptive awareness, thought suppression, perceived stress, and academic success*.

Megı́as-Robles, A., Gutiérrez-Cobo, M. J., Gómez-Leal, R., Cabello, R., Gross, J. J., & Fernández-Berrocal, P. (2019). Emotionally intelligent people reappraise rather than suppress their emotions. *PloS One*, *14*(8), e0220688.

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.

Sai, L., Shang, S., Tay, C., Liu, X., Sheng, T., Fu, G., Ding, X. P., & Lee, K. (2021). Theory of mind, executive function, and lying in children: A meta-analysis. *Developmental Science*, *24*(5), e13096.

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

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.

Sip, K. E., Carmel, D., Marchant, J. L., Li, J., Petrovic, P., Roepstorff, A., McGregor, W. B., & Frith, C. D. (2013). When pinocchio’s nose does not grow: Belief regarding lie-detectability modulates production of deception. *Frontiers in Human Neuroscience*, *7*, 16.

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.

Stewart, S. L., Wright, C., & Atherton, C. (2019). Deception detection and truth detection are dependent on different cognitive and emotional traits: An investigation of emotional intelligence, theory of mind, and attention. *Personality and Social Psychology Bulletin*, *45*(5), 794–807.

Ten Brinke, L., Lee, J. J., & Carney, D. R. (2019). Different physiological reactions when observing lies versus truths: Initial evidence and an intervention to enhance accuracy. *Journal of Personality and Social Psychology*, *117*(3), 560.

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.

Veer, A. E. van’t, Stel, M., Beest, I. van, & Gallucci, M. (2014). Registered report: Measuring unconscious deception detection by skin temperature. *Frontiers in Psychology*, *5*, 442.

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

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