**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; **leal2008time?** ), resulting in physiological arousal (Eskritt et al., 2021) which may require effortful self-regulation to conceal(Krokoszinski & Hosser, 2016). While much of deception research has focused on topics such as theory of mind, other mechanisms underlying the regulation of physiological tells, such as increased cardiac arousal, respiration and sweating that are often used for lie detection tests have been under researched.

Interoception refers to a cluster of skills involved in monitoring, interpreting, integrating, and regulating one’s visceral states(Chen et al., 2021; Weiss et al., 2014), and is closely tied to emotional and cognitive processes (Barrett & Simmons, 2015; Petzschner et al., 2021). Garfinkel et al. (2015) consolidated the facets of interoception into three distinct processes: *interoceptive accuracy*, as an objective measure of one’s ability to monitor viscera; *interoceptive sensibility*, as the subjective measure of accuracy; and *interoceptive awareness*, describing the metacognitive perception of interoceptive accuracy. While interoceptive sensibility may be an important component, the corrospondence between the confidence and accuracy, or interoceptive awareness appears to be associated with the downregulation of emotional responses and their associated physiological symptoms (Critchley & Garfinkel, 2017, 2018; Füstös et al., 2013).

Interoceptive processes are not only a key component forming your own emotions, but also indicates being in tune with the emotional states of others (Garfinkel et al., 2016; Konvalinka et al., 2011). This corresponds to situations involving deception, since much of the existing literature examining the link between interoceptive abilities has focused on how it may implicitly flag that one is being lied to through an elevated cardiac response (Gunderson et al., 2021), and skin temperature (Veer et al., 2014). While research indicates that attending to ones interoceptive cues improves success in unmasking detection(Ten Brinke et al., 2019), there is limited information how interoceptive skills may function the other way around, in the case of the liar.

As identified by Murphy (2022), the propensity to use interoceptive information, and in what manner remains an underdeveloped area in research, and may illuminate the mechanisms underlying emotional control exhibited by successful liars (Krokoszinski & Hosser, 2016; Wielgopolan & Imbir, 2021).

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.

# Methods

Along with open science standards, the material (stimuli generation code, experiment code, raw data, analysis script with complementary figures and analyses, preregistration, etc.) for this research is available at: **[LINK].**

The plan for this study was preregistered **(OSF?)**.

## Participants

*26* participants (65.4% Females, M*age* = 20.9, *SD* = 2.0) were recruited advertising using univerisy mailing lists and posters. Participants were awarded with academic credit for their participation. This research was approved by the **Institutional Review Board (Reference Number: IRB-)** of Nanyang Technological University (NTU).

## 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.

### 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).

### 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. During the first phase of the experiment, demographic measures were taken, followed by measures of Theory of Mind and interoceptive abilities, as described above. Participants were then connected to an electrocardiogram before proceeding with the deception task on a computer.

#### Deception Task.

Participants underwent 40 trails of a directed lying task, where they were asked questions such as, “What secondary school did you go to?”. Answers were provided verbally and participant pressed a button when they were finished, at which point they would be connected for 10 seconds to the condition-unique stimulus. There were two conditions of 20 trails each, beginning with the *interrogation* *condition*, and followed by the second, *polygraph condition*. The purpose of these conditions was to draw out the interoceptive or theory of mind abilities by providing feedback cues that would cause them to focus their attention on either the visual feedback cues from other people, or their own cardiac response.

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. We hypothesized that in this condition, participants would rely on their theory of mind abilities to discern the facial feedback cues from the examiner to ascertain whether their lie was successful.

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 hypothesised that when participants were made acutely aware of such signals through false feedback, they 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.

Debriefing was provided following the study. At this time the truthful answers to the directed lie questions were collected via questionnaire.

## 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 modeled 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, particularly 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 underwent two conditions of false feedback where they were intended to rely on either interoceptive or theory of mind abilities. To our knowledge, this is the first study which has shown a clear link between interoceptive abilities and cardiac regulation during the act of deception. As this relationship was largely found in the polygraph condition that was designed to draw out the interoceptive abilities within the individual, we can conclude that our manipulation was successful. This taken with high confidence ratings in participant’s lying abilities, also indicates other potential factors at play.

Our results show that participants who were telling the truth had a slower reaction time, and rated themselves as less confident in the polygraph condition, however, their heart rate averaged lower as compared to the interrogation condition. This suggests that despite the participant’s lack of confidence, they were very in-tune with their interoceptive feedback in the polygraph condition, and successfully able to decrease their heart rate. Given that interoceptive accuracy is enhanced by focusing on it during a time of stress (Schulz et al., 2013), such as attempting to conceal deception combined with false physiological feedback delivered during the polygraph condition, this conclusion is in line with our hypothesis.

As the relationship between interoception and emotion is well established (Barrett & Simmons, 2015; Dobrushina et al., 2020; Ohira, 2020), our work is consistent with Wielgopolan and Imbir (2021), who found that groups with higher emotional awareness made better liars and more successful at regulating their emotional and physiological responses, as well as fabricating emotion.

Although participants were not explicitly directed to attend to their physiological signals as a means of evading deception detection, they were aware these signals were recorded and informed. Additionally, the presence of the ECG likely made them aware, and these measures remained connected during the interrogation. This indicates that there was likely an explicit, top-down attentional component to the control exerted, which is further reflected by a broader theme in our results. Our results indicate that all interoceptive factors (*Meta, Listening, Focus and Regulation*) are associated with an increased confidence in deceptive abilities, and faster reaction time for lies in the polygraph condition. It logical that confidence in deception would be accompanied by faster reaction time. However, since confidence was a subjective measure, it does not necessarily reflect competence (Khalsa et al., 2008). While ourdata indicates that participants were able to self-regulate using interoceptive abilities, this still highlights an important caveat, which is the potential metacognitive awareness at play (Garfinkel et al., 2015).

Metacognitive beliefs, or interoceptive awareness about competence in the quality of physiological regulation may influence the cognitive processes and development physiological responses. That is to say, that the appraisal is just as important to the process of deception by informing their true ability (Harrison et al., 2021; Petzschner et al., 2021; Stephan et al., 2016).

High interoceptive sensibility regardless of true competence may also directly shape behaviour through the reduced perception of risk associated with being caught. Past research indicates the likelihood of deception increases when individuals judge a lower risk of being caught, based on situation-specific variables and their own qualities (Sip et al., 2012, 2013). These results may offer support to findings such as Vabba et al. (2022) where interoceptive abilities were found to be a significant moderator in the willingness to engage in riskier forms of deception, and offer insight into other risk behaviours linked with interoceptive ability such as cheating (Williams et al., 2016) and risk-taking (Ditto et al., 2006; Lenggenhager et al., 2013).

While this research has demonstrated that interoceptive abilities may be used to regulate and conceal deception, the individual’s confidence in their lying abilities, is subjective, and may not offer a full-picture of their objective success. Future research should continue to unpack this relationship, examining how confidence and interoceptive regulation is reflected in the actual success of a lie, as well as how individuals confidence in their interoceptive abilities may correspond to decision bias (Garfinkel et al., 2015), and propensity to engage in deception (Sip et al., 2012). Furthermore, linking and expanding recent research (Pearson & Pfeifer, 2022) on how individual traits correlate with true interoceptive abilities, in addition to interoceptive sensibility as they relate to 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.

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

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.

Critchley, H. D., & Garfinkel, S. N. (2017). Interoception and emotion. *Current Opinion in Psychology*, *17*, 7–14.

Critchley, H. D., & Garfinkel, S. N. (2018). The influence of physiological signals on cognition. *Current Opinion in Behavioral Sciences*, *19*, 13–18.

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.

Dobrushina, O., Dobrynina, L., Arina, G., Kremneva, E., Suslina, A., Gubanova, M., Belopasova, A., Solodchik, P., Urazgildeeva, G., & Krotenkova, M. (2020). Interaction of interoceptive perception and emotional intelligence: A functional neuroimaging study. *Neuroscience and Behavioral Physiology*, *50*(8), 1043–1050.

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.

Garfinkel, S. N., Tiley, C., O’Keeffe, S., Harrison, N. A., Seth, A. K., & Critchley, H. D. (2016). Discrepancies between dimensions of interoception in autism: Implications for emotion and anxiety. *Biological Psychology*, *114*, 117–126.

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*.

Harrison, O. K., Garfinkel, S. N., Marlow, L., Finnegan, S. L., Marino, S., Köchli, L., Allen, M., Finnemann, J., Keur-Huizinga, L., Harrison, S. J., et al. (2021). The filter detection task for measurement of breathing-related interoception and metacognition. *Biological Psychology*, *165*, 108185.

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

Khalsa, S. S., Rudrauf, D., Damasio, A. R., Davidson, R. J., Lutz, A., & Tranel, D. (2008). Interoceptive awareness in experienced meditators. *Psychophysiology*, *45*(4), 671–677.

Konvalinka, I., Xygalatas, D., Bulbulia, J., Schjødt, U., Jegindø, E.-M., Wallot, S., Van Orden, G., & Roepstorff, A. (2011). Synchronized arousal between performers and related spectators in a fire-walking ritual. *Proceedings of the National Academy of Sciences*, *108*(20), 8514–8519.

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

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.

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.

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.

Murphy, J. (2022). Propensity to use interoceptive signals: An important individual difference. *Biological Psychology*, 108326–108326.

Ohira, H. (2020). Predictive processing of interoception, decision-making, and allostasis: A computational framework and implications for emotional intelligence. *Psihologijske Teme*, *29*(1), 1–16.

Pearson, A., & Pfeifer, G. (2022). Two measures of interoceptive sensibility and the relationship with introversion and neuroticism in an adult population. *Psychological Reports*, *125*(1), 565–587.

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.

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

Schulz, A., Lass-Hennemann, J., Sütterlin, S., Schächinger, H., & Vögele, C. (2013). Cold pressor stress induces opposite effects on cardioceptive accuracy dependent on assessment paradigm. *Biological Psychology*, *93*(1), 167–174.

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.

Stephan, K. E., Manjaly, Z. M., Mathys, C. D., Weber, L. A., Paliwal, S., Gard, T., Tittgemeyer, M., Fleming, S. M., Haker, H., Seth, A. K., et al. (2016). Allostatic self-efficacy: A metacognitive theory of dyshomeostasis-induced fatigue and depression. *Frontiers in Human Neuroscience*, *10*, 550.

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.

Wielgopolan, A., & Imbir, K. K. (2021). Can emotional awareness of liars influence deception detection effectiveness? *Journal of Personality Assessment*, *103*(5), 667–674.

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