UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Semantic Leakage Enables Lie Detection, but First-Person Pronouns and Verbosity Can Get in the Way of Detection

Permalink
https://escholarship.org/uc/item/83z6r9jk

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors
Yeter, Ōzlem
Kooi, Barteld P.
de Weerd, Harmen
et al.

Publication Date
2024

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Semantic Leakage Enables Lie Detection, but First-Person Pronouns and Verbosity Can Get in the Way of Detection

Özlem Yeter[1] (o.yeter@rug.nl)
Barteld Kooi[1] (b.p.kooi@rug.nl)
Harmen de Weerd[2] (harmen.de.weerd@rug.nl)
Rineke Verbrugge[2] (l.c.verbrugge@rug.nl)
Petra Hendriks[3] (p.hendriks@rug.nl)

[1]Department of Theoretical Philosophy, University of Groningen, The Netherlands
[2]Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence, University of Groningen, The Netherlands
[3]Center for Language and Cognition Groningen, University of Groningen, The Netherlands

Abstract

We investigated the impact of linguistic cues and autistic traits on lie detection. Adult participants (N = 125) judged suspects' statements in a detective game. Untruthful statements were marked by semantic leakage. Literature indicates that liars use fewer first-person pronouns and mental-state terms than truth-tellers. We manipulated the untruthful statements for the presence/absence of these cues to test their effect on lie detection. The adults were 89% accurate in detecting lies. Mental-state terms did not affect accuracy, while presence of first-person pronouns hindered it. Having autistic traits did not influence lie detection. However, adults with higher autistic traits struggled to detect lies when these contained both a first-person pronoun and a mental-state term. Post-hoc analysis revealed lower lie detection accuracy for longer sentences. Our findings underscore the significance of semantic leakage in lie detection, with nuanced effects of linguistic cues on accuracy, particularly for adults with higher autistic traits.

Keywords: lie detection; semantic leakage; linguistic cues; autistic traits**;** theory of mind

Introduction

Being able to distinguish truth from lies is crucial for one's safety and healthy relationships. This ability is particularly crucial given the risks (e.g., manipulation and abuse) that susceptibility to lies poses for people who experience difficulties with social communication, such as people with autism spectrum disorder (ASD). Nevertheless, some meta-analyses covering over a hundred studies on the ability to detect lies suggest that adults demonstrate accuracy levels comparable to chance in experimental settings (e.g., Aamodt & Custer, 2006; Bond & DePaulo, 2006). One of the primary reasons for this chance-level accuracy could be people's assumption that conversations are in general cooperative, in particular, that speakers obey Grice's maxim of quality (Grice, 1975). People often assume that what they are told is true, because questioning everything would make it challenging to effectively engage in conversation. This tendency is further supported by studies that argue for a truth bias, suggesting that individuals choose to believe an interlocutor unless presented with a reason not to (Levine, 2014; Levine, 2022). In contrast, police officers and detectives, who have more encounters with liars when they are required to investigate a case, are slightly better than the general population at lie detection (Ekman O'Sullivan, & Frank 1999; Mann, Vrij, Bull, & 2004). Therefore, a key question is: How do people detect lies?

Semantic Leakage

Liars may develop successful linguistic strategies to form a lie, starting from early childhood (Hu et al., 2020). In addition to employing effective verbal tactics, liars must also inhibit telling the truth and be cautious not to reveal anything that could put them on the spot while lying (Evans & Lee, 2013; O'Connor, Dykstra, & Evans, 2020). After all, there is little room for doubt when liars reveal themselves by leaking critical information (Brennen & Magnussen, 2022). Suppose, for example, that a detective asks a suspect: "Jack was found dead three days ago. Are you involved in his death?" Now suppose that the suspect answers "I didn't kill Jack; it's terrible that he was shot to death". In this case, the detective can be almost certain that the suspect is lying, because the suspect provides specific and unprompted information – in this case, mentioning the cause of death, '*shot to death'*. It is impossible for the suspect to know this without having been informed or involved. This phenomenon of providing information that one could not know if one were speaking the truth is called *semantic leakage*.

The majority of studies that focused on semantic leakage and lies investigated semantic leakage control during the *production* of lies in children (Lavoie & Talwar, 2020; Talwar & Lee, 2002; Talwar & Lee, 2008). However, to the best of our knowledge, no study has addressed the concept of semantic leakage from the perspective of the *receiver* of a lie.

Linguistic Cues to Lie Detection

Researchers from diverse fields have delved into various methods of detecting lies. While in the past the polygraph was a popular technique for detecting lies by monitoring subjects' physiological activity, its reliability remains questionable (Iacono & Ben-Shakhar, 2019). To date, there is no fool-proof method or technique to detect lies. Recently, the focus has shifted from physiological cues to linguistic cues in lie production (Ditmarsch, Hendriks, & Verbrugge, 2020).

Newman and colleagues (2003) discovered that untruthful statements contained more motion verbs than mental-state

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.* ©2024 The Author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

verbs (e.g., go, take, walk rather than know, believe, feel), in comparison to truthful statements. They argued that this preference is due to the cognitive simplicity and accessibility of motion verbs compared to mental-state verbs. They added that since liars' cognitive resources are occupied with fabricating lies, liars involuntarily avoid exhausting their cognitive resources by using fewer mental-state verbs.

Another pattern that has been documented in untruthful statements is less use of first-person pronouns than in truthful statements (Hancock et al., 2007; Newman et al., 2003). This might be due to the latent need to dissociate oneself from the act of deception (Knapp, Hart, & Dennis, 1974; Walczyk et al., 2013).

The Role of Theory of Mind

In addition to these linguistic cues, an important factor in lie detection could be theory of mind (ToM) (Premack & Woodruff, 1978). Previous studies pointed to ToM as a crucial mechanism in the process of lie detection in childhood (Maas, 2010; Nancarrow et al., 2018). ToM, a sub-component of social cognition, enables people to navigate social communication by inferring the mental states (beliefs, emotions, intentions, desires, etc.) of others and recognising that their own mental states might differ from those of others. ToM may be relevant in lie detection, because one needs to assess the intentions, beliefs, and motivations of a possible liar. In an experiment with adults, Stewart, Wright and Atherton (2019) highlight the role of ToM as a building block of lie detection.

A particular group that is often found to struggle with ToM, experiencing difficulty in attributing mental states to others as well as in perspective taking, are people with ASD (Baron-Cohen et al., 1995; Bellesi et al., 2016; Senju, 2012). Some studies proposed that individuals with ASD have difficulty in detecting deception due to their impaired ToM ability, making them more susceptible to manipulation (Fisher, Moskowitz, & Hodapp, 2013; Williams et al., 2018).

The Present Study

This study aims to contribute to the literature by investigating adults' ability to detect lies in the presence of semantic leakage. In particular, we investigated: i) the effect of linguistic cues, namely, the presence of first-person pronouns and mental-state terms, on lie detection accuracy; and ii) the relationship between having autistic traits and the ability to detect lies. Based on the preceding literature review, we expect the following:

1. Semantic leakage will be taken as a reliable cue for lie detection.
2. Linguistic cues, namely, absence of first-person pronouns and absence of mental-state terms, will facilitate lie detection.
3. Higher autistic traits will hinder lie detection.

Method

The experiment consisted of three parts: a questionnaire asking some background characteristics, an autism questionnaire assessing autistic traits, and a Detective Game testing participants' ability to detect lies.

Participants

The data were collected online using Qualtrics. In total, 125 adults (61 female) between the ages of 18 and 65 (mean age: 30.76) participated in our study. Each participant was given a compensation of 8 Euros for their participation. To motivate the participants to answer as accurately and carefully as possible, after answering all questions, a random question from the Detective Game was automatically selected for each participant. If the participant had answered that question correctly, they were awarded an additional 2 Euros (see Allais, 1953; Azrieli, Chambers, & Healy, 2018). The participants were told about this procedure at the start of the experiment. The study was reviewed by the Ethical Review Committee CETO of the University of Groningen. Each participant gave informed consent prior to their participation.

Design and Materials

The Detective Game The experiment testing adults' ability to detect lies was presented as a detective game. It included 16 short stories (8 baseline stories, 8 test stories) about a crime or transgression and a suspect of this crime or transgression. The critical statements in each test story were two sentences uttered by the suspect. The second critical statement uttered by the suspect contained semantic leakage. In the first critical statement, two linguistic cues were manipulated in a 2x2 design: presence of first-person pronoun (+/-1PP) and presence of mental-state term (+/-MST). Thus, there were 4 conditions, named A, B, C, D. Below is a test story from the experiment, illustrating all four:

Excerpted story 1:
Jane feels a soft touch on her bag as she is walking in a crowded street. She reaches for her bag to check it and realizes that her wallet is missing. She immediately looks back and sees a woman standing close to her. Jane suspects her and asks:
"Did you take my wallet?"

The woman replies as follows:

Condition A: *"Nobody would steal anything in this crowd. Especially not for so little money."* (-1PP, -MST)

Condition B: "*Nobody would think of stealing anything in this crowd. Especially not for so little money."* (-1PP, +MST)

Condition C: *"I have never stolen anything. Especially not for so little money."* (+1PP, -MST)

Condition D: *"I would not think of stealing anything. Especially not for so little money*." (+1PP, +MST)

For baseline stories, in contrast, the sentences uttered by the suspect did not contain any semantic leakage and did not contain the same manipulation of linguistic cues as the test stories (4 critical statements had both a first-person pronoun and a mental-state term, while the other 4 contained neither). They were intended to represent truthful statements. Below is a baseline story from the experiment:

Excerpted story 2:
Emma is a museum guard in a big city. On one of her duty days, there is a power cut in the museum building. The whole building stays in the dark for 12 minutes after the power goes off. The next day, Emma realises that one of the vases in the museum is gone. The museum officers identify the visitors that entered the museum on that day and time. They interrogate one of the visitors and ask the following:
"A vase went missing from the museum yesterday. Are you involved in this act?"
The visitor states the following:
"It was clear that something was wrong when the power was gone for so long. But I came to the museum without a bag and did not move at all during the power cut."

To avoid any biases due to specific properties of the stories, we created four conditions for each test story and distributed them across four lists in a counterbalanced manner. Each list consisted of 8 baseline stories and 8 test stories (2 test stories per condition). Participants were assigned to one of these lists randomly. The four lists contained the same test stories but in different conditions. The order of presentation of the baseline and test stories was randomized within the list.

After reading the critical statements (i.e., suspects' statements), participants were asked whether they thought the suspect was guilty of the crime or transgression or not. A *Guilty* judgement meant that the participant found the statement untruthful. A *Not guilty* judgement indicated that the participant found the statement truthful. If the participant believed the suspect to be guilty, they were asked to justify their answers. The target response was *Guilty* for test stories and *Not guilty* for baseline stories.

Autism-Spectrum Quotient Questionnaire The Autism-Spectrum Quotient (AQ; Baron-Cohen Wheelwright, Skinner, Martin, & Clubley 2001) is a self-report measure used to investigate autistic traits in people aged 16 and older. The AQ has demonstrated satisfactory test–retest reliability (r = .70) and consistent internal reliability, with Cronbach's alpha scores ranging from .63 to .77 (cf. Austin, 2005; Hoekstra, Bartels, Cath, & Boomsma, 2008). The AQ consists of 50 questions using a 4-way Likert scale. About half of the items were worded to prompt a "slightly/definitely disagree" response, while the other half aimed to elicit a "slightly/definitely agree" response from individuals with ASD. Answers that indicate having autistic traits received a score of 1 point, while the rest received a score of 0. The questions fall into 5 categories, with 10 questions per category: social skill, attention switching, attention to detail, communication, and imagination. Following Baron-Cohen and colleagues' (2001) scoring method, we looked at total scores on the AQ, which ranged from 0 points (no autistic traits) to 50 points (full agreement on all autistic traits statements). A threshold of 26 is taken as an indicator for a potential case for ASD.

Background Questionnaire Participants' background characteristics were collected through a questionnaire asking about their age, gender, experiences with lying, and whether they had a formal diagnosis of ASD.

Procedure

Before the actual experiment, participants saw two practice items, on which they received immediate feedback (correct, incorrect). They did not receive feedback on items in the actual experiment.

There were two instructional manipulation check items, which checked for participants' attention during the experiment (see Oppenheimer, Meyvis, & Davidenko, 2009). Participants had to answer both attention check questions correctly for their data to be included in the analysis. One participant's data was not included in the analysis due to failing this attention check. The other inclusion criterion was having more than beginner level of English proficiency, in order to avoid potential challenges in fully comprehending the stories and justifying the responses in English. No participant reported to have beginner level of English proficiency.

Analysis of the Detective Game

Accuracy rates per story were inspected to check whether the stories themselves had an effect on participants' accuracy. The accuracy rates of test story nr. 7 were much lower (mean: 50.33%) than the other stories (mean: 89.47%). Upon closer inspection, one of the critical statements of this story was ambiguous and hence the semantic leakage may not have been clear to participants. Therefore, this test story was not included in our analyses, leaving 7 test stories to be analysed.

The statistical analysis was conducted using SPSS. First, test items and baseline items were compared for responses (i.e., *Guilty*, *Not guilty*). Subsequent analyses were done only for the test items. Using univariate analysis, first, participants' accuracy levels across the four test conditions A, B, C, D (see example story above) were investigated. Next, binary logistic regression analysis was run to explore the effect of linguistic cues (absence/presence of first-person pronouns and mental-state terms) on lie detection accuracy. Our dependent variable consisted of participants' responses on test items (scored as 0 or 1). Subsequently, we investigated the relationship between participant characteristics (age, English proficiency, AQ scores) and lie detection accuracy using linear regression analysis, with the accumulative score (ranging from 0 to 7) for the test items as the dependent variable. Finally, the same analysis was done to investigate the effect of lying and lie detection habits on the accumulative lie detection accuracy score.

Results

Background Characteristics

Participant background characteristics are presented in Table 1. Of the 125 participants, 4 reported having been diagnosed with ASD.

Table 1: Participant responses on 5-point Likert-scale background questionnaire.

	M	SD	SE
English proficiency	3.93	.541	.048
Detecting others' lies [1]	3.01	.893	.080
Others detecting your lies [1]	2.88	.989	.088
Telling other-benefiting lies [2]	2.59	1.29	.115
Telling self-benefiting lies [2]	2.46	1.059	.095

Lie Detection Accuracy

There were 1875 items in total in the analysis (875 test items, 1000 baseline items). Figure 1 shows participants' responses on baseline items compared to test items. Test items (i.e., items with semantic leakage) are judged as *Guilty* more often (N = 782 versus N = 183) compared to baseline items (i.e., items without semantic leakage). Comparing the percentages of *Guilty* judgements between baseline items and test items (89.37% versus 18.30%) using paired samples t-test, the difference was significant [t(124) = -31.345, SD = 25.35 $p < .001$].

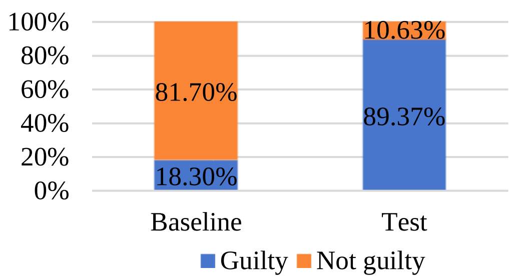

Figure 1: Distribution of participant responses on baseline items vs. test items.

Comparing participants' responses on the four test conditions (A, B, C, D), participants had the lowest accuracy for condition C, where the critical statements contained a first-person pronoun but no mental-state verb (see Figure 2).

We ran a binary logistic regression with the two linguistic cues (+/-1PP or +/-MST in critical statements) and their interactions as predictor variables for lie detection accuracy. Absence of first-person pronouns was the only significant predictor for better lie detection accuracy in this model: When there was no first-person pronoun, participants gave more correct (*Guilty*) judgements compared to when a first-person pronoun was present (B = -.639, SE = .304, Wald χ^2 = 4.420, p = .036, 95% CI [-75.56, -66.58]).

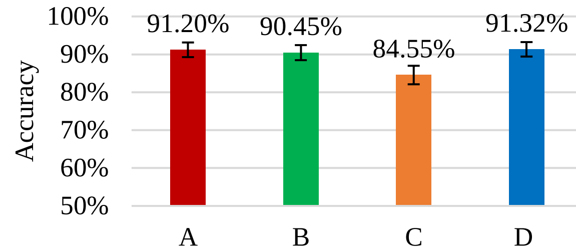

Figure 2: Lie detection accuracy on test items across conditions. Linguistic cues per condition: A = -1PP, -MST; B = -1PP, +MST; C = +1PP, -MST; D = +1PP, +MST.

In analysing participants' justifications, we found that they noticed semantic leakage at a minimum rate of 90% of the cases regardless of test condition (A, B, C, D). Two examples of participant justifications in response to excerpted story 1 (see Design and Materials) are:

- *She must have taken it to know that it has so little money*.
- *The woman mentioned the amount of the money although it is not mentioned by Jane.*

Baseline (truthful) items consisted of two conditions. Half (4) of the items consisted of +1PP, +MST sentences. In 5% of these items participants commented on the indirectness and evasiveness of suspects. For judgements in the other 4 items, with -1PP, -MST, this indirectness or evasiveness was mentioned in 11% of the items. Nevertheless, in both conditions, participants wrongly judged sentences as untruthful almost at the same rate (~19%). This indicates that sentences that lack both first-person pronouns and mental-state terms are more easily taken to be indirect and evasive, but this does not cause these sentences to be taken to be more untruthful than the +1PP, +MST sentences that contain both of them. Overall, this finding highlights the significant effect of semantic leakage as a cue for lie detection.

Autism Spectrum Quotient Scores and Participant Characteristics

The mean total AQ score from a possible range of 0-50 was 19.99 (SD = 7.33, Range = 6 – 41). A Kolmogorov-Smirnov test supported the assumption of normality [D(125) = .066, p = 0.200]. Figure 3 displays the histogram for the total AQ score. Comparing the total AQ scores of the four participants who reported to have received a formal diagnosis of ASD (M = 32.75, SD = 5.91) to the rest of the sample (M =19.57, SD = 7.03), the four participants with diagnosed ASD had significantly higher AQ scores [F(1, 123) = 13.70, $p < .001$]. This confirms the appropriateness of the AQ questionnaire as a tool to measure autistic traits. For the subscales of AQ, the

[1] level of easiness (from extremely difficult to extremely easy)

[2] level of frequency (from never/rarely to every day)

mean score was 3.62 (SD = 2.30) for *social skill*, 5.11 (SD = .21) for *attention switching*, 5.32 (SD = .22) for *attention to detail*, 2.82 (SD = .20) for *communication*, 3.13 (SD = .169) for *imagination*.

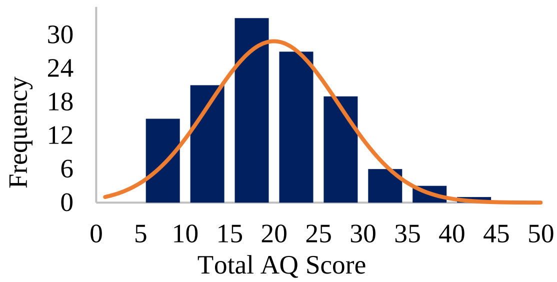

Figure 3: Distribution of participants' AQ scores.

Next, we ran a linear regression analysis to examine the effect of participant characteristics on lie detection accuracy. There were 3 independent variables: age, English proficiency, and total AQ score; our dependent variable was lie detection accuracy. ANOVA for this model failed to explain the variance in lie detection accuracy [$F(3,121) = .657$, $p = .580$]. These independent variables did not predict the accuracy levels in truth detection for baseline items either ($p = .318$). Subsequently, we conducted another regression analysis replacing the total AQ score with five AQ subscales, namely, social skill, attention switching, attention to detail, communication, and imagination. This model did not yield a statistically significant prediction for lie detection accuracy either [$F(7, 117) = 1.206$, $p = .305$].

In the background questionnaire, we had inquired about participants' frequency of lying for their own benefit and for other people's benefit, as well as about their self-perceived ability to detect other people's lies and how easily others could detect their lies (see Table 1). The answers that participants provided on the four questions in the background questionnaire were taken as predictors for lie detection. According to the linear regression analysis, there was no predictor effect of lying and lie detection habits on lie detection accuracy [$F(4,120) = .716$, $p = .582$].

Post-hoc Analyses

To follow up on our analyses, given that the critical statements in the four conditions differed in length (see Figure 4), we were interested in investigating whether the length of critical statements had any effect on lie detection accuracy, in addition to the linguistic cues. Since the manipulation of the linguistic cues occurred in the first part of the critical statements, and the length of the second part was the same across conditions, we calculated the word count for the first part only. We introduced this word count as an additional variable, *length of utterance*, in our logistic regression analysis in addition to the two linguistic cues (+/- 1PP, +/- MST). Both first-person pronoun (B = -.657, SE = .306, Wald $\chi^2 = 4.595$, $p = .032$, Exp(B) = .518, 95% CI [.284, .945]) and length of utterance (B = -.179, SE = .058, Wald $\chi^2 = 9.616$, $p = .002$, Exp(B) = .836, 95% CI [.747, .936]) were significant predictors for lie detection accuracy, both in a negative direction: Participants were more likely to correctly judge the critical statements in the test items as a lie when this statement did not contain a first-person pronoun and when it was shorter.

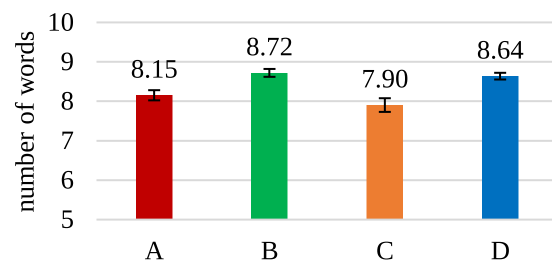

Figure 4: Length of utterances on test items across conditions. Linguistic cues per condition: A = -1PP, -MST; B = -1PP, +MST; C = +1PP, -MST; D = +1PP, +MST.

There is literature suggesting that individuals with ASD face difficulties in comprehending texts that contain mental-state terms (Mathersul, McDonald, & Rushby, 2013). Another study conducted by Mizuno and colleagues (2011) showed that processing first-person pronouns is more challenging for adults with ASD compared to neurotypical adults. Considering that test condition D contained both a mental-state term and a first-person pronoun, we were interested in running an additional post-hoc analysis to investigate whether lie detection accuracy for condition D is predicted by AQ scores. We ran a binary logistic regression analysis because we entered lie detection accuracy as a binary (0 or 1) dependent variable. We created a dummy variable representing condition D and put this variable, AQ scores and their main interactions as independent variables in this model. The analysis predicted lie detection accuracy as a function of AQ score and condition D. For condition D, lower AQ scores were associated with higher lie detection accuracy [B = -.075, SE = .036, Wald $\chi^2 = 4.258$, $p = .039$, Exp(B) = .928, 95% CI (.747, .936)]. Considering that there is no significant main effect for AQ, these results suggest that participants that scored high on the AQ questionnaire encountered difficulties in judging statements in condition D (+1PP, +MST), while accuracies on conditions A, B, and C seem to be independent of the AQ score.

Discussion

The majority of the participants in this study were highly accurate in both detecting lies (89%) and identifying sentences without a lie (81%). This outstanding performance presents a contrasting picture when compared to the literature, which often emphasizes chance-level accuracy in adults' lie detection (e.g., Ekman et al., 1999; Hancock et al., 2007). Semantic leakage appears to be a reliable cue that helps people in detecting lies.

The primary objective of this study was to examine the impact of linguistic cues on adults' accuracy in detecting scripted lies in the presence of semantic leakage, and whether having autistic traits has a detrimental effect on accuracy. Our findings provide strong evidence that adults perceive statements with semantic leakage (indicating a lie) accurately by judging the suspect as *Guilty*. To investigate the impact of linguistic cues on lie detection, our focus centred on two key elements derived from the existing literature on lie production patterns: fewer first-person pronouns and fewer mental-state terms. In other words, we tested whether lie receivers could accurately pick up on the linguistic pattern of lies described in the literature. We tested this by manipulating these cues in statements that contained a lie marked by semantic leakage. The presence of first-person pronouns and the use of more words in these statements led several participants to incorrectly judge the suspects in our experiment as *Not guilty* (truthful), despite the presence of semantic leakage acting as counterevidence.

Autistic traits assessed through AQ (Baron-Cohen et al., 2001) did not form a significant factor in overall lie detection accuracy. However, when statements with semantic leakage contained both a first-person pronoun and a mental-state term, participants with high AQ scores detect lies at a lower rate than participants with low AQ scores. The remainder of this section will discuss these findings in further detail.

Participants in this study incorrectly perceived the baseline statements as lies by judging the suspects as *Guilty* about 19% of the time. This was a higher rate than the *Not guilty* judgements (11% on average) in test items that contained a lie. This shows that the truth bias (as described by Levine, 2014) is not universal, at least it seems absent in the context of a ‘detective game’ where participants are tasked to be suspicious.

Hancock and colleagues (2007) found that liars produced less self-oriented first-person pronouns in their untruthful statements to divert the attention from themselves. However, our results suggest that this strategy backfires: Statements containing a first-person pronoun were perceived as more truthful than those that lack it. We further refer to the justifications of participants who correctly judged the suspects as *Guilty* in the test conditions that lacked a first-person pronoun. Some of these participants highlighted the indirectness of the suspect’s statement and their attempt to put the blame off themselves. Using language lacking first-person pronouns seems to cause a sense of indirectness and evasiveness. Nevertheless, when we compare baseline stories containing different linguistic cues, we do not see any difference between participants’ accuracy rates. This suggests that relying solely on linguistic cues is not always sufficient to assess the truthfulness of a sentence. In contrast, semantic leakage provides a strong cue for detecting lies, even when the statements include first-person pronouns.

Van Swol, Braun and Malhotra (2012) demonstrated that people use more words when lying, which they call the “Pinocchio effect”. Furthermore, they found that even truthful participants were inaccurately seen as deceptive when they used linguistic cues associated with lying. Our findings paint a somewhat different picture, because participants in this study tended to incorrectly judge long utterances as truthful compared to shorter ones, which does match with liars’ patterns in the literature. Although we did not investigate the cause for this, it could be that longer sentences are perceived as too detailed to be lies (for a review, see DePaulo et al., 2003). Alternatively, the attention of lie receivers may be focused more on comprehending detailed information than on lie detection.

Unlike Williams and colleagues (2018), we did not observe a significant effect of having autistic traits on overall lie detection accuracy. Studies investigating the relationship between ASD and ToM in an adult context provide mixed results (Mathersul et al., 2013; cf. Gernsbacher & Frymiare, 2005; Tager-Flusberg & Sullivan, 1994). In their review, Gernsbacher and Yergeau (2019) highlight the predictor role of language abilities in successful ToM tasks and suggest that the link between ASD and ToM skills should be interpreted carefully. Considering that language development improves with age, obtaining results similar to studies on children might not happen when testing adult participants. Moreover, only 4 participants in this study had previously been diagnosed with autism, which did not allow us to see the effect of having an ASD diagnosis on lie detection.

For participants with higher autistic traits, lie detection accuracy was lower for statements that contained both a first-person pronoun and a mental-state term. This is in line with the literature suggesting that people with ASD have difficulty processing first-person pronouns and mental-state terms (Mathersul et al., 2013; Mizuno et al., 2011). Inclusion of these linguistic cues may make the process of lie detection, which is already a cognitively demanding task, more difficult for people who have high autistic traits.

While this study yields a novel contribution to the existing literature by being the first to investigate the concept of semantic leakage from the perspective of lie receivers, it is not without limitations. We instructed participants to provide justifications when identifying a story suspect as *Guilty*. However, a more comprehensive understanding of why participants considered sentences with semantic leakage trustworthy could have been attained if we had also asked them to provide justifications when finding the story suspects *Not guilty*. However, we opted against this to keep the experiment under an hour in length. Also, we could not control the length of utterances across stories, because we needed to manipulate linguistic cues. While this variation allowed us to observe the significant effect of verbosity on lie detection, keeping length constant could have enhanced our investigation of linguistic cues (first-person pronouns and mental-state terms) in the lie detection process.

To conclude, we propose that lie detection is highly accurate in the presence of semantic leakage. However, the presence of first-person pronouns and verbosity generally decrease accuracy, and the combined presence of first-person pronouns and mental-state terms specifically decreases accuracy for individuals with higher autistic traits.

References

Aamodt, M. G., & Custer, H. (2006). Who can best catch a liar ? A meta-analysis of individual differences in detecting deception. *The Forensic Examiner*.

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école Américaine. *Econometrica*, *21*, 503–546. https://doi.org/10.2307/1907921

Austin, E. J. (2005). Personality correlates of the broader autism phenotype as assessed by the Autism Spectrum Quotient (AQ). *Personality and Individual Differences*, *38*(2), 451–460. https://doi.org/10.1016/j.paid.2004.04.022

Azrieli, Y., Chambers, C. P., & Healy, P. J. (2018). Incentives in experiments: A theoretical analysis. *Journal of Political Economy*, *126*(4), 1472–1503.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, *31*(1), 5–17.

Baron-Cohen, S., Campbell, R., Karmiloff-Smith, A., Grant, J., & Walker, J. (1995). Are children with autism blind to the mentalistic significance of the eyes? *British Journal of Developmental Psychology*, *13*(4), 379–398. https://doi.org/10.1111/j.2044-835x.1995.tb00687.x

Bellesi, G., Jameel, L., Vyas, K., Crawford, S., & Channon, S. (2016). Using and reasoning about social strategies in autism spectrum disorder in everyday situations. *Research in Autism Spectrum Disorders*, *25*, 112–121. https://doi.org/10.1016/j.rasd.2016.02.004

Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review*, *10*(3), 214–234. https://doi.org/10.1207/s15327957pspr1003_2

Brennen, T., & Magnussen, S. (2022). The science of lie detection by verbal cues: What are the prospects for its practical applicability? *Frontiers in Psychology*, *13*(835285). https://doi.org/10.3389/fpsyg.2022.835285

DePaulo, B. M., Malone, B. E., Lindsay, J. J., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin*, *129*(1), 74–118. https://doi.org/10.1037/0033-2909.129.1.74

Ditmarsch, H., Hendriks, P., & Verbrugge, R. (2020). Editors' review and introduction: Lying in logic, language, and cognition. *Topics in Cognitive Science*, *12*(2), 466–484. https://doi.org/10.1111/tops.12492

Ekman, P., O'Sullivan, M., & Frank, M. G. (1999). A few can catch a liar. *Psychological Science*, *10*(3), 263–266. https://doi.org/10.1111/1467-9280.00147

Evans, A. D., & Lee, K. (2013). Emergence of lying in very young children. *Developmental Psychology*, *49*(10), 1958–1963. https://doi.org/10.1037/a0031409

Fisher, M. H., Moskowitz, A. L., & Hodapp, R. M. (2013). Differences in social vulnerability among individuals with autism spectrum disorder, Williams syndrome, and Down syndrome. *Research in Autism Spectrum Disorders*, *7*(8), 931–937. https://doi.org/10.1016/j.rasd.2013.04.009

Gernsbacher, M. A., & Frymiare, J. L. (2005). Does the autistic brain lack core modules? *The Journal of Developmental and Learning Disorders*, *9*, 3–16. http://www.ncbi.nlm.nih.gov/pubmed/25520587

Gernsbacher, M. A., & Yergeau, M. (2019). Empirical failures of the claim that autistic people lack a theory of mind. *Archives of Scientific Psychology*, *7*(1), 102–118. https://doi.org/10.1037/arc0000067

Grice, H. P. (1975). Logic and conversation. In *Syntax and Semantics 3: Speech Acts.* https://doi.org/http://dx.doi.org/10.1016/j.eiar.2014.08.006

Hancock, J. T., Curry, L. E., Goorha, S., & Woodworth, M. (2007). On lying and being lied to: A linguistic analysis of deception in computer-mediated communication. *Discourse Processes*, *45*(1), 1–23. https://doi.org/10.1080/01638530701739181

Hoekstra, R. A., Bartels, M., Cath, D. C., & Boomsma, D. I. (2008). Factor structure, reliability and criterion validity of the Autism-Spectrum Quotient (AQ): A study in Dutch population and patient groups. *Journal of Autism and Developmental Disorders*, *38*(8), 1555–1566. https://doi.org/10.1007/s10803-008-0538-x

Hu, C., Huang, J., Wang, Q., Weare, E., & Fu, G. (2020). Truthful but misleading: Advanced linguistic strategies for lying among children. *Frontiers in Psychology*, *11*(April). https://doi.org/10.3389/fpsyg.2020.00676

Iacono, W. G., & Ben-Shakhar, G. (2019). Current status of forensic lie detection with the comparison question technique: An update of the 2003 National Academy of Sciences report on polygraph testing. *Law and Human Behavior*, *43*(1), 86–98. https://doi.org/10.1037/lhb0000307

Knapp, M. L., Hart, R. P., & Dennis, H. S. (1974). An exploration of deception as communication construct. *Human Communication Research*, *1*(1), 15–29. https://doi.org/10.1111/j.1468-2958.1974.tb00250.x

Lavoie, J., & Talwar, V. (2020). Care to Share? Children's Cognitive Skills and Concealing Responses to a Parent. *Topics in Cognitive Science*, *12*(2), 485–503. https://doi.org/10.1111/tops.12390

Levine, T. R. (2014). Truth-Default Theory (TDT): A theory of human deception and deception detection. *Journal of Language and Social Psychology*, *33*(4), 378–392. https://doi.org/10.1177/0261927X14535916

Levine, T. R. (2022). Truth-default theory and the psychology of lying and deception detection. *Current Opinion in Psychology*, *47*, 101380. https://doi.org/10.1016/j.copsyc.2022.101380

Maas, F. K. (2010). Children's understanding of promising, lying, and false belief. *The Journal of General Psychology*, *135*(3), 301–322.

https://doi.org/10.3200/GENP.135.3.301-322

Mann, S., Vrij, A., & Bull, R. (2004). Detecting true lies: Police officers' ability to detect suspects' lies. *Journal of Applied Psychology*, *89*(1), 137–149. https://doi.org/10.1037/0021-9010.89.1.137

Mathersul, D., McDonald, S., & Rushby, J. A. (2013). Understanding advanced theory of mind and empathy in high-functioning adults with autism spectrum disorder. *Journal of Clinical and Experimental Neuropsychology*, *35*(6), 655–668. https://doi.org/10.1080/13803395.2013.809700

Mizuno, A., Liu, Y., Williams, D. L., Keller, T. A., Minshew, N. J., & Just, M. A. (2011). The neural basis of deictic shifting in linguistic perspective-taking in high-functioning autism. *Brain*, *134*(8), 2422–2435. https://doi.org/10.1093/brain/awr151

Nancarrow, A. F., Gilpin, A. T., Thibodeau, R. B., & Farrell, C. B. (2018). Knowing what others know: Linking deception detection, emotion knowledge, and theory of mind in preschool. *Infant and Child Development*, *27*(5), 1–9. https://doi.org/10.1002/icd.2097

Newman, M. L., Pennebaker, J. W., Berry, D. S., & Richards, J. M. (2003). Lying words: Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin*, *29*(5), 665–675. https://doi.org/10.1177/0146167203029005010

O'Connor, A. M., Dykstra, V. W., & Evans, A. D. (2020). Executive functions and young children's lie-telling and lie maintenance. *Developmental Psychology*, *56*(7), 1278–1289. https://doi.org/10.1037/dev0000955

Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology*, *45*(4), 867–872. https://doi.org/10.1016/j.jesp.2009.03.009

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*(4), 515–526. https://doi.org/10.1017/S0140525X00076512

Stewart, S. L. K., 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. https://doi.org/10.1177/0146167218796795

Tager-Flusberg, H., & Sullivan, K. (1994). A second look at second-order belief attribution in autism. *Journal of Autism and Developmental Disorders*, *24*(5), 577–586. https://doi.org/10.1007/BF02172139

Talwar, V., & Lee, K. (2002). Development of lying to conceal a transgression: Children's control of expressive behaviour during verbal deception. *International Journal of Behavioral Development*, *26*(5), 436–444. https://doi.org/10.1080/01650250143000373

Talwar, V., & Lee, K. (2008). Social and cognitive correlates of children's lying behavior. *Child Development*, *79*(4), 866–881. https://doi.org/10.1111/j.1467-8624.2008.01164.x

Van Swol, L. M., Braun, M. T., & Malhotra, D. (2012). Evidence for the Pinocchio effect: Linguistic differences between lies, deception by omissions, and truths. *Discourse Processes*, *49*(2), 79–106. https://doi.org/10.1080/0163853X.2011.633331

Walczyk, J. J., Igou, F. P., Dixon, A. P., & Tcholakian, T. (2013). Advancing lie detection by inducing cognitive load on liars: A review of relevant theories and techniques guided by lessons from polygraph-based approaches. *Frontiers in Psychology*, *4*(February), 1–13. https://doi.org/10.3389/fpsyg.2013.00014

Williams, D. M., Nicholson, T., Grainger, C., Lind, S. E., & Carruthers, P. (2018). Can you spot a liar? Deception, mindreading, and the case of autism spectrum disorder. *Autism Research*, *11*(8), 1129–1137. https://doi.org/10.1002/aur.1962