

Running head: Small Talk and Comparable Truth Baselines

Detecting Deception Through Small Talk and Comparable Truth Baselines

Nicola Palena¹, Letizia Caso¹, Aldert Vrij², and Robin Orthey²

¹ Department of Human and Social Sciences, University of Bergamo

² Department of Psychology, University of Portsmouth

Correspondence concerning this article should be addressed to: Nicola Palena, Department of Human and Social Sciences, University of Bergamo. Piazzale Sant'Agostino, 2, 24129 Bergamo (Italy).

Email: n.palena8376a@ordpsicologier.it

Authors note. We would like to thank Guido Carlotto, Monica Marciali and Lara De Mizio for their help with data collection

Abstract

The present experiment investigates similarities in participants' nonverbal and verbal behaviours when responding to baseline and investigative questions, comparing two different types of baselines. Police literature suggests to obtain a baseline through small talk, whereas academic literature underlines the importance of baseline and investigative themes to be comparable. First, a baseline was obtained (either small talk or comparable), then the investigative questioning started. During the investigative questioning participants either truthfully reported a set of actions they had actually performed or lied about them. Findings revealed that truth tellers and liars in the small talk condition did not differ in their level of similarity when responding to the baseline and investigative questions. In the comparable truth condition, levels of verbal similarity between the baseline and investigative questions were higher for truth tellers than for liars but only for one variable: spatial detail. Results therefore showed that a small talk baseline should not be used to assess interviewees' credibility, and that a comparable truth baseline, although better than a small talk baseline, is still problematic.

Keywords: detecting deception, baseline, similarity scores

Detecting Deception Through Small Talk and Comparable Truth Baselines

Meta-analyses into (non)verbal cues to deception have shown that such cues are faint and unreliable (DePaulo, Lindsay, Malone, Muhlenbruck, Charlton & Cooper, 2003; Hartwig & Bond, 2011). Consequently, researchers started to focus on specific interview techniques to enhance existing or elicit new cues to deceit (Vrij & Granhag, 2012). These include cognitive lie detection (Vrij, 2014, 2015; Vrij, Fisher, Blank, Leal & Mann, 2016), the Assessment Criteria Indicative of Deception (Colwell, Hiscock-Anisman, Memon, Taylor & Prewett, 2007; Colwell, Hiscock-Anisman, Memon, Colwell, Taylor & Woods, 2009), and the Strategic Use of Evidence (Granhag & Hartwig, 2008, 2015).

These techniques have been developed within the academic setting. On the contrary, a strategy frequently proposed and used within the police culture is the use of a small talk baseline (Ewens, Vrij, Jang & Jo, 2014; Frank, Yarbrough & Ekman, 2006; Inbau, Reid, Buckley & Jayne, 2013). The core of the baseline technique is that interviewers should in first instance evaluate interviewees' baseline behaviour using non-threatening questions (small talk) and then assess if differences in behaviours between this phase and the investigative phase of the interview arise. If this is the case, then it is thought that deception may be occurring.

Frank and colleagues (2006) proposed such an implementation of the baseline technique, called the Improving Interpersonal Evaluations for Law Enforcements and National Security Technique (Frank *et al.*, 2006). Others believe that this version of the technique is misleading (Ewens *et al.*, 2014, Moston & Engelberg, 1993). The problem is that behaviours in response to non-threatening questions cannot be compared to behaviours in response to questions asked in the investigative interview. The baseline component is a relatively low-stakes situation whereas the interview is a relatively high-stakes situation and people show different behaviours in low versus high stakes situations (Hartwig & Bond, 2014; Ioannou & Hammond, 2015). In fact, behaviours are influenced by many factors other than deception. People react differently when they are accused of

Running head: Small Talk and Comparable Truth Baselines

wrongdoing than when they are unchallenged (Vrij, 2006), and show different behaviours when they are interviewed by different people (Vrij & Winkel, 1991). Moreover, behaviour is topic-related. People respond differently when discussing a topic that embarrasses them than when discussing a neutral topic (Kleinke, 1986), and they respond differently when they discuss a topic they care about or is important to them than when they discuss a topic with which they have less personal involvement (Davis & Hadiks, 1995; Matarazzo, Wiens, Jackson, & Manaugh, 1970). Finally, people's behaviour sometimes changes over time in the same interview (Buller & Burgoon, 1996; Burgoon, Buller, White, Afifi, & Buslig, 1999; Stiff, Corman, Krizek, & Snider, 1994), or, if they are interviewed on more than one occasion, changes may occur over repeated interviews (Granhag & Strömwall, 2002).

Therefore, when researchers wish to compare a person's deceptive nonverbal response with a truthful nonverbal response from the same person, they need to make sure that the deceptive and truthful responses are taken from the same interview setting; that the person talks about similar topics in the deceptive and truthful parts; and that these parts were discussed within a short period of time from each other. In that respect, Vrij (2008) and Ewens *et al.* (2014) refer to a *comparable truth* baseline. Comparable means that the baseline the investigator uses must be similar in content, stakes, and cognitive and emotional involvement to investigative questions. Additionally, the contexts within which the baseline and the target questioning occur must be similar (Vrij, 2008).

In the present experiment we compared the effect of a comparable truth baseline with the effect of a small talk baseline, typically used by police and proposed by Frank *et al.* (2006). We compared a series of nonverbal and verbal behaviours displayed during baseline and the interview and calculated a similarity score: The higher the score, the higher the similarity. We did not expect the similarity scores for nonverbal or verbal behaviours to differ between truth tellers and liars in the small talk baseline condition, due to problems associated with this type of baseline comparison outlined above. The comparable truth baseline technique might work, and perhaps even more so for

Running head: Small Talk and Comparable Truth Baselines

verbal behaviours, as these are typically more diagnostic cues to deceit than nonverbal behaviours (Bond & DePaulo, 2006). We therefore predicted the following interaction effect: Truth tellers and liars will obtain similar similarity scores in the small talk baseline condition but truth tellers will obtain higher similarity scores in the comparable truth baseline condition (Hypothesis 1). The effect sizes in the comparable truth baseline condition may be more pronounced for verbal behaviours than for nonverbal behaviours (Hypothesis 2).

Method

Participants

A total of 69 participants (49 female subjects and 20 male subjects) took part in the experiment. Age ranged from 21 to 53 years, with an average age of $M = 26.88$ years old ($SD = 8.10$).

Design

The experiment utilised a 2 (Baseline: comparable truth vs. small talk) X 2 (Veracity: truth vs. lie) between subjects design. Participants were all honest in Phase one (baseline) but they were randomly assigned to the type of baseline and veracity conditions. Regarding the type of baseline, participants either had to report truthfully a set of actions they were asked to perform which were similar to the set of actions they had to perform and were interviewed about in the actual interview (comparable truth condition) or had to tell something truthfully about their last year as a student and/or as a worker (small talk condition). Regarding the veracity condition, participants had either to tell the truth or to lie about a second set of task they were asked to perform. The dependent variables were the similarity scores for hands and finger movements, one arm movements, two arms movements, spatial, temporal, visual, audio and action details. We examined hands and fingers movements as meta-analyses and literature reviews show that it can discriminate between truth tellers and liars to a moderate extent (DePaulo *et al.*, 2003; Vrij, 2008). We examined one arm and two arms movements as physical rigidity may be related to attempted control, which is thought to

Running head: Small Talk and Comparable Truth Baselines

be related to lying (Zuckerman, DePaulo & Rosenthal, 1981). Finally, we examined spatial, temporal, visual, audio and action details as they are related to the verbal veracity tool Reality Monitoring, which has a solid theoretical foundation (Johnson & Raye, 1981; Vrij, 2008). The level of motivation, measured on a 5 points likert scale, was set as a covariate.

Procedure

A list of participants was obtained during university lectures. Students who accepted to participate were offered two additional points for a university exam, but only if they could convince the interviewer that they were telling the truth. Eventually, all participants received the two points. They were also informed that the experiment lasted approximately 35-40 minutes. Upon arrival, the participants completed an informed consent form. Then, the experimenter told them that they were going to perform a set of tasks and that after completion of these tasks they would be interviewed. They were also informed that the interviews would be video-recorded. The set of tasks included two sub-sets of tasks. For the first sub-set of tasks, the participants received an envelope from the experimenter, which contained the password needed to log onto a laptop. The experimenter then left the room. Once logged onto the laptop the participants had to read the only word document file that was available on the PC desktop. This word document instructed participants to search for a backpack in the room, which contained a CD-ROM and to watch the video file from the CD-ROM. They could play, pause, stop and view the video as many time as they needed. On the video a man appeared in view who informed the participants that they had to search for a key inside the same backpack where they found the CD-ROM. That key could open a safe-deposit box that was placed near the window. The man on the video also informed the participants that they would find further instructions inside the safe-deposit box. These instructions informed participants that they had to access the Internet, send an email to a specific email address and to exit the room and wait for a person (a confederate).

The meeting with the confederate served as a landmark. Everything that happened before the

Running head: Small Talk and Comparable Truth Baselines

meeting was part of the comparable truth baseline event, whereas everything that happened after the meeting was part of the target event. However, only participants in the comparable truth condition had to report what they had done *before* the meeting (in addition to what they had done after the meeting).

Regarding the second sub-set of actions, the confederate gave the participants a newspaper, which contained further instructions and left the room. The instructions informed the participants that they had to enter the room adjacent to where they performed the first set of tasks and to look for a coat hook on which an USB stick was located. Participants had to take the USB stick and to look for a book inside the wardrobe in the room in which they would find another USB stick. They had to switch the two USB sticks and to keep the new one for the remaining part of the experiment. They were also asked to leave the newspaper near the book inside the wardrobe. The participants were then asked to return to the first room and to wait for the interviewer.

A 56 years old male psychology student, who was blind to both Baseline and Veracity experimental conditions, acted as the interviewer. When he arrived he introduced himself and asked an open-ended question regarding either the first sub-set of tasks (comparable truth baseline condition) or the last year of study/work (small talk baseline condition). Then the interviewer left the room and the confederate returned to give instructions about the veracity condition: Participants allocated to the truthful condition were told to continue to answer the questions truthfully. Participants allocated to the lying condition were instructed to lie to the questions that would be asked regarding the second sub-set of tasks. They were instructed to create a detailed and credible story, which included at least 4-5 actions or events. When liars confirmed that they understood the instructions, the confederate left the room and told them that they had five minutes to prepare the lie. Truth tellers were also given five minutes to prepare themselves. After five minutes the interviewer came back into the room and asked a second open-ended question regarding what had happened after the meeting. After the interview participants completed a questionnaire investigating

Running head: Small Talk and Comparable Truth Baselines

background characteristics and their motivation to do well in the interview (on a 5-point Likert scale ranging from [1] low to [5] high). Finally, participants were debriefed.

Coding

A coder blind to the hypotheses and the experimental conditions coded the interviews for the frequency per minute of hands and finger movements (movements of hands and fingers while arms are held still), one arm movements (movements of one arm with the other remaining still) and two arms movements (movements of both arms at the same time). The coder further coded the frequency for spatial (information about place and/or spatial arrangement of objects and people), temporal (information about when an event happens and or the description of a sequence of events), visual (information about what the interviewee saw: “I saw a man entering the room”), audio (information about what the interviewee heard: “I heard a men shouting”) and action details (“I wrote an email). A second coder, again blind to hypotheses and experimental conditions, coded 18 interviews (25%) for the same verbal and nonverbal behaviours to evaluate the inter-rater agreement (Interclass Correlation Coefficient). The lowest level of agreement (ICC = .75) was obtained for the variable *Visual details*, whereas the highest level of agreement was obtained for the variable *Actions* and was of (ICC = .96). Therefore, the scores indicated high inter-rater agreement.

Similarity scores computation

The similarity scores were computed as follows: for each dependent variable, we divided the lowest score by the highest score, regardless of the phase of the interview. This means that if the lowest score for a variable appeared in the baseline, we divided the score of that variable in the baseline by the score of the same variable in the target phase. The same was true for the opposite situation. Then, we multiplied the result by 100. The highest the score, the more similar participants behaved between the two phases of the interview.

Results

To assess whether the level of motivation differed between conditions, an ANOVA was

Running head: Small Talk and Comparable Truth Baselines

conducted with Baseline (comparable truth vs. small talk) and Veracity (truth tellers vs. liars) as factors and the level of motivation as the dependent variable. A main effect for Baseline appeared, $F(1, 65) = 8.40, p < .01$, Cohen's $d = .69$, with participants in the comparable truth condition showing higher level of motivation ($M = 4.46, SD = .56$) than those in the small talk condition ($M = 3.97, SD = .82$). The main effect for Veracity, $F(1, 65) = .80, p = .78$ and the Baseline X Veracity interaction effect, $F(1, 65) = .01, p = .91$ were not significant.

To test our hypotheses, first, a MANOVA for nonverbal cues was conducted with Baseline (comparable truth vs. small talk) and Veracity (truth tellers vs. liars) as the between-subjects factors and the similarity scores for hands and fingers, one arm and two arm movements as dependent variables. At a multivariate level, the main effect for Baseline, $F(3, 63) = 6.15, p = .001$ was significant. The main effect for Veracity, $F(3, 63) = .22, p = .88$, and the Baseline X Veracity interaction, $F(3, 63) = 2.22, p = .09$ were not significant.

At a univariate level, the main effect for Baseline was significant for one arm movements only, $F(1, 65) = 18.27, p < .001$. Participants in the comparable truth baseline ($M = 68.16, SD = 26.62, C. I. [59.29; 77.04]$) showed more similarity than participants in the small talk baseline ($M = 38.45, SD = 30.72, C. I. [27.37; 49.53]$). On the other hand, participants in the comparable truth baseline obtained similar similarity scores for hands and finger movements ($M = 71.91, SD = 23.40, C. I. [64.11; 79.71]$) to participants in the small talk baseline ($M = 64.78, SD = 28.10, C. I. [54.65; 74.91]$), $F(1, 65) = 1.15, p = .29$. Likewise, those in the comparable truth baseline ($M = 53.78, SD = 37.00, C. I. [41.45; 66.12]$) did not differ from those in the small talk baseline ($M = 60.06, SD = 36.55, C. I. [46.89; 73.24]$) for what concerns two hands movements, $F(1, 65) = .48, p = .49$.

Results, therefore, did not support Hypothesis 1 for nonverbal cues, as the interactions effect was not significant. However, the fact that participants in the comparable truth baseline obtained higher similarity scores for one arm movements than those in the small talk baseline, supports the idea that interviewees interviewed via a small talk baseline tend to change their behaviour more than those

Running head: Small Talk and Comparable Truth Baselines

interviewed via a comparable truth baseline.

Second, a MANCOVA for verbal cues was conducted with Baseline (comparable truth vs. small talk) and Veracity (truth tellers vs. liars) as the between-subjects factors and the similarity scores for spatial, temporal, visual, audio and action details as the dependent variables.

Additionally, since these details may vary with the overall number of details within each of the interview phases (Baseline and Target), the total number of details provided in the Baseline and the Target phases were introduced as covariates.

At a multivariate level, the main effect for Baseline, $F(5, 59) = 12.02, p < .001$, and the Baseline X Veracity interaction, $F(5, 59) = 2.52, p = .04$, were significant. The Veracity main effect was not significant, $F(5, 59) = .57, p = .72$.

At a univariate level, there was a main effect for Baseline for spatial, $F(1, 63) = 11.35, p = .001$, and visual, $F(1, 63) = 51.03, p < .001$ similarity scores. The only significant Baseline X Veracity effect appeared for similarity scores for spatial details, $F(1, 63) = 9.24, p = .003$. Since the interaction effect is the most informative effect and considering that our hypotheses are related to it, only the interaction effect will be discussed. We were particularly interested in differences between truth tellers and liars in the two baseline conditions. To be as informative as possible, we report the data for all dependent variables.

Comparable truth condition

A MANOVA was conducted on the comparable truth condition data with Veracity (truth tellers vs. liars) as factor. The dependent variables were the similarity scores for hands and fingers, one arm and two arms movements. At a multivariate level, the Veracity effect was not significant, $F(3, 33) = , p = .43$. At a univariate level, no significant effect appeared (Table 1).

A MANCOVA was conducted on the same data with Veracity as factor, spatial, temporal, visual, audio, and action similarity scores as the dependent variables and the total amount of details in the two interview phases as the covariates. At a multivariate level, the veracity effect was not

Running head: Small Talk and Comparable Truth Baselines

significant, $F(5, 29) = 1.74, p = .15$. At a univariate level, a significant effect for spatial detail similarity scores appeared. Truth tellers ($M = 78.67, SD = 15.60, C. I. [71.14; 86.18]$) showed more similarity than liars ($M = 52.25, SD = 26.75, C. I. [38.95; 65.55]$), $F(1, 33) = 7.44, p = .01$, Cohen's $d = 1.20$.

All univariate effects and their effect sizes are reported in Table 1. Three out of eight variables showed similarity scores that went in the 'wrong' direction (less similarity amongst truth tellers), but for those five who went in the predicted direction the strongest effect size was for spatial details, $d = 1.20$, followed by hands and finger movements, $d = .54$. The effect sizes for the remaining three variables were small (d - scores ranged between .10 and .25). This means that Hypothesis 2, the effect sizes would be stronger for verbal behaviours than for nonverbal behaviours, only obtained very limited support.

Enter Table 1 about here

To examine whether the similarity scores differed from perfect similarity (100) one-sample t-tests were conducted for each dependent variable and were tested against a 100 score. If the comparable truth works, effects should be significant for liars only. The results are reported in Table 2, which shows that all effects were significant, for both truth tellers and liars. The high d -scores imply that the responses from both liars and truth tellers in the interview differed considerably from their responses in the comparable truth baseline.

Enter Table 2 about here

Small talk condition

A MANOVA was conducted on the small talk condition data with Veracity (truth tellers vs. liars) as factor. The dependent variables were the similarity scores for hands and fingers, one arm and two arms movements. The multivariate effect for Veracity was not significant, $F(3, 28) = 1.43, p = .25$. None of the univariate effects was significant either (Table 3).

A MANCOVA was conducted with the same data again with Veracity (truth tellers vs. liars)

Running head: Small Talk and Comparable Truth Baselines

as factor and spatial, temporal, visual, audio, and action similarity scores as dependent variables. The total number of details provided in the Baseline and the Target phases were set as covariates. At a multivariate level, the Veracity effect was not significant, $F(5, 24) = .57, p = .72$. No significant univariate effect appeared neither (Table 3).

In sum, the multivariate effects were not significant in the comparable truth nor in the small talk conditions. In the comparable truth condition truth tellers displayed greater similarity on one variable compared to no significant difference at all in the small talk condition. This means that the comparable truth baseline was more effective than the small talk baseline, supporting Hypothesis 1, but even in the comparable truth condition the findings were weak.

Similar to the comparable truth condition analyses, a one-sample t-test was conducted for each dependent variable and were tested against a 100 score. If a small talk baseline is ineffective, all results should be significant. Table 2 shows that this was indeed the case.

Enter Table 3 about here

Discussion

In this experiment, we tested the efficacy of two different types of baselines for deception detection using a similarity rating. We found that similarity scores did not differ between truth tellers and liars for participants in the small talk condition. This result is similar to that obtained by Ewens and colleagues (2014), who found that both truth tellers and liars showed differences in the amount of hard thinking and behavioural control between two phases of the interview. Our results therefore reject the efficacy of a small talk baseline procedure as proposed by Frank and colleagues (Frank *et al.*, 2006) and by Inbau and colleagues (Inbau *et al.*, 2013) as in such a procedure truth tellers and liars would appear equally deceptive.

We found a significant difference between truth tellers and liars in the comparable truth condition, with truth tellers showing more similarity than liars. However, it is too premature to conclude that a comparable truth baseline works for at least two reasons. First, the method only

Running head: Small Talk and Comparable Truth Baselines

worked for one of the variable examined in the experiment: spatial details. It is possible that this result is due to the type of task. Indeed, the set of actions participants discussed in the comparable baseline phase and actual interview phase showed similarity in spatial information (moving around two rooms and to interact with several objects placed in several places). Truth tellers showed more spatial similarity in discussing these two sets of spatial activities than liars, who told the truth about one set of spatial activities but lied about the other set of activities. Thus, to create an effective baseline the type of activities discussed in the baseline and target interview need to be comparable. Second, even for the spatial detail variable the findings in the comparable truth condition were problematic because, although truth tellers showed more similarity than liars, they did not show perfect similarity. This means that in real life, judging whether or not a difference in spatial similarity between baseline and target interview occurred will not work to detect deceit. Somehow, the size of dissimilarity needs to be taken into account, which will be difficult to do.

Furthermore, we employed a procedure whereby the veracity instructions were given after the baseline and before the target phase of the interview. We opted for this strategy as there was the risk of liars not following the instruction to be as truthful as possible during the baseline phase. That is, informing them about their veracity condition could have influenced how they answered to the baseline questions (Deeb *et al.*, 2017; Strömwall & Willén, 2011). Had this been the case, we would have not had a real and truthful baseline. Perhaps participants in the present study thought about this strategy when they were given time to prepare themselves for the target phase of the interview. That is, once the interviewee was given the instruction to lie after the baseline phase, s/he may have employed consciously the strategy to try to provide responses in the target question that were as similar as possible to those given in the baseline question. Future studies could explore how the comparable truth baseline technique works when the veracity instruction are given before the baseline part of the interview.

Further research could also explore the baseline technique when no preparation time is given

Running head: Small Talk and Comparable Truth Baselines

between the two phases. Although our results are interesting there is the risk that our interview was perceived as two separated interviews rather than a single one. We do not think this was the case, the time lapse we provided was enough to prepare but in our opinion not too long to make the interview look like two different interviews, but there is a small risk that at least some participants considered this to be two interviews.

Limitations

The stakes for this experiment were relatively low. However, as Ewens *et al.* (2014) noted, higher stakes are unlikely to make a small talk baseline more effective. Rather, this should further decrease the efficacy of a small talk baseline procedure as differences in stakes between baseline and investigative questions would become even more pronounced. Raising the stakes should not be an issue for the comparable truth baseline approach as one of the assumptions is to keep the stakes of the two phases constant so that they remain comparable.

Additionally, our sample size was small. Therefore, although we found some interesting results, further research should explore how the similarity scoring system performs when bigger samples are used.

Conclusion

Despite being promoted in the academic and professional literature, a baseline lie detection technique using a small talk baseline is not effective in discriminating truth tellers from liars as similar differences arise for both Veracity groups. Practitioners should stay away from using such a baseline procedure. The alternative baseline procedure, the comparable baseline procedure, was more successful than the small talk baseline procedure but still not good enough to be implemented in real life. There are two options. First, drop the baselining technique altogether. Given the paucity of support for this technique, nothing would be lost. Second, because baselining is popular amongst practitioners, researchers could spend more time designing an appropriate baseline method. This article revealed some difficulties researchers will face when designing an appropriate method.

References

- Bond, C. F. & DePaulo, B. M. (2006) Accuracy of deception judgments. *Personality and Social Psychology Review, 10*(3), 214-234.
- Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory, 6*, 203-242. DOI: 10.1111/j.1468-2885.1996.tb00127.x
- Burgoon, J. K., Buller, D. B., White, C. H., Afifi, W., & Buslig, A. L. (1999). The role of conversational involvement in deceptive interpersonal interactions. *Personality and Social Psychology Bulletin, 25*, 669-686. DOI: 10.1177/0146167299025006003
- Colwell, K., Hiscock-Anisman, C. K., Memon, A., Colwell, L. H., Taylor, L., & Woods, D. (2009). Training in Assessment Criteria Indicative of Deception (ACID) to improve credibility assessment. *Journal of Forensic Psychology Practice, 9*, 199-207.
- Colwell, K., Hiscock-Anisman, C. K., Memon, A., Taylor, L., & Prewett, J. (2007). Assessment Criteria Indicative of Deception (ACID): An integrated system of investigative interviewing and detecting deception. *Journal of Investigative Psychology and Offender Profiling, 4*(3), 167-180. DOI: 10.1002/jip.73
- Davis, M., & Hadiks, D. (1995). Demeanor and credibility. *Semiotica, 106*, 5–54. DOI: 10.1515/semi.1995.106
- Deeb, H., Vrij, A., Hope, L., Mann, S., Granhag, P. A., & Lancaster, G. L. (2017). Suspects' consistency in statements concerning two events when different question formats are used. *Journal of Investigative Psychology and Offender Profiling, 14*(1), 74-87.
- DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin, 129*(1), 74-118.
- Ewens, S., Vrij, A., Jang, M., & Jo, E. (2014). Drop the small talk when establishing baseline behaviour in interviews. *Journal of Investigative Psychology and Offender Profiling*. DOI: 10.1002/jip.1414

Running head: Small Talk and Comparable Truth Baselines

- Frank, M. G., Yarbrough, J. D., & Ekman, P. (2006). Investigative interviewing and the detection of deception. In T. Williamson (Eds.), *Investigative interviewing: Rights, research and regulation* (pp. 229–255). Cullompton, Devon: Willan.
- Granhag, P. A., & Hartwig, M. (2008). A new theoretical perspective on deception detection: On the psychology of instrumental mind reading. *Psychology, Crime & Law, 14*(3), 189-200.
- Granhag, P. A., & Hartwig, M. (2015). The Strategic Use of Evidence technique: A conceptual overview. In P. A. Granhag, A. Vrij, & B. Verschuere (Ed), *Detecting deception: Current challenges and cognitive approaches* (pp. 231-252). Chichester: John Wiley and Sons.
- Granhag, P. A., & Strömwall, L. A. (2002). Repeated interrogations: Verbal and non-verbal cues to deception. *Applied Cognitive Psychology, 16*, 243-257. DOI: 10.1002/acp.784
- Hartwig, M., & Bond Jr., C. F. (2011). Why do lie-catchers fail? A lens model meta-analysis of human lie judgments. *Psychological Bulletin, 137*(4), 643-659. DOI: <http://dx.doi.org/10.1037/a0023589>
- Hartwig, M., & Bond Jr., C. F. (2014). Lie Detection from Multiple Cues: A Meta-analysis. *Applied Cognitive Psychology, 28*, 661–676. DOI: 10.1002/acp.3052
- Inbau, F. E., Reid, J. E., Buckley, J. P., & Jayne, B. C. (2013). *Criminal interrogation and confessions*, 5th edition. Burlington, MA: Jones & Bartlett Learning.
- Ioannou, M., & Hammond, L. (2015). The detection of deception within investigative contexts: Key challenges and core issues. *Journal of Investigative Psychology and Offender Profiling, 12*(2), 107-118. DOI: 10.1002/jip.1433
- Kleinke, C. L. (1986). Gaze and eye contact: a research review. *Psychological bulletin, 100*, 78. DOI: 10.1037/0033-2909.100.1.78
- Matarazzo, J. D., Wiens, A. N., Jackson, R. H., & Manaugh, T. S. (1970). Interviewee speech behavior under conditions of endogenously-present and exogenously-induced motivational states. *Journal of Clinical Psychology, 54*, 15-26. DOI: 10.1037/h0028651

Running head: Small Talk and Comparable Truth Baselines

- Moston, S., & Engelberg, T. (1993). Police questioning techniques in tape-recorded interviews with criminal suspects. *Policing and Society: An International Journal of Research and Policy*, *3*, 223-237. DOI: 10.1080/10439463.1993.9964670
- Stiff, J., Corman, S., Krizek, B., & Snider, E. (1994). Individual differences and changes in nonverbal behavior unmasking the changing faces of deception. *Communication Research*, *21*, 555-581. DOI: 10.1177/009365094021005001
- Strömwall, L. A., & Willén, R. M. (2011). Inside criminal minds: Offenders' strategies when lying. *Journal of Investigative Psychology and Offender Profiling*, *8*(3), 271-281.
- Vrij, A. (2006). Challenging interviewees during interviews: The potential effect on lie detection. *Psychology, Crime & Law*, *12*, 193-206.
- Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities*. Chichester: John Wiley and Sons.
- Vrij, A. (2014). Interviewing to detect deception. *European Psychologist*, *19*, 184-194. DOI: 10.1027/1016-9040/a000201
- Vrij, A. (2015). A cognitive approach to lie detection. In P. A. Granhag, A. Vrij, & B. Verschuere (Ed), *Detecting deception: Current challenges and cognitive approaches* (pp. 205-230). Chichester: John Wiley and Sons.
- Vrij, A., Fisher, R. P., Blank, H., Leal, S., & Mann, S. (2016). A cognitive approach to elicit verbal and nonverbal cues to deceit. In J-W. van Prooijen, & P. A. M. van Lange (Eds.), *Cheating, corruption, and concealment: the roots of dishonesty*. Cambridge: Cambridge University Press.
- Vrij, A., & Granhag, P. A. (2012). Eliciting cues to deception and truth: What matters are the questions asked. *Journal of Applied Research in Memory and Cognition*, *1*(2), 110-117. DOI: 10.1016/j.jarmac.2012.02.004
- Vrij, A., & Winkel, F. W. (1991). Cultural patterns in Dutch and Surinam nonverbal behavior: An

Running head: Small Talk and Comparable Truth Baselines

analysis of simulated police/citizen encounters. *Journal of Nonverbal Behavior*, 15, 169–
184. DOI: 10.1007/BF01672219

Table 1. *Similarity scores as a function of Veracity in the comparable truth condition*

Similarity scores	Truth Tellers		Liars		<i>F</i>	<i>p</i>	<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
Hands and fingers movements	77.95	23.53	65.54	22.12	2.72	.10	.54
One arm movements	69.52	27.82	66.73	26.02	.09	.75	.10
Two arms movements	52.87	39.42	54.74	35.38	.02	.88	-.05
Spatial details	78.67	15.60	52.25	26.75	7.44	.01	1.20
Temporal details	43.51	33.14	55.79	30.18	1.09	.30	-.38
Visual details	70.56	17.95	65.53	21.12	.53	.47	.25
Audio details	14.47	32.61	30.46	39.78	.51	.48	-.44
Action details	66.09	18.21	61.96	22.89	.00	.98	.20

Table 2. *Statistics examining whether the similarity scores differ from perfect similarity*

	Comparable Truth				Small Talk			
	Truth Tellers		Liars		Truth Tellers		Liars	
	<i>t</i>	<i>d</i>	<i>t</i>	<i>d</i>	<i>t</i>	<i>d</i>	<i>t</i>	<i>d</i>
Spatial details	-5.961	-1.40	-7.574	-1.83	-11.425	-2.85	-6.285	-1.68
Temporal details	-7.429	-1.75	-6.213	-1.50	-6.309	-1.57	-6.081	-1.62
Visual details	-7.147	-1.68	-6.922	-1.69	-28.682	-7.17	-12.890	-3.44
Audio details	-11.431	-2.69	-7.415	-1.80	-4.781	-1.19	-5.909	-1.58
Action details	-8.115	-1.91	-7.048	-1.71	-10.115	-2.53	-6.501	-1.74
Hands and fingers movements	-4.084	-.96	-6.609	-1.60	-6.265	-1.56	-4.027	-1.11
One arm movements	-4.775	-1.12	-5.424	-1.31	-7.497	-1.87	-10.607	-2.94
Two arms movements	-5.210	-1.23	-5.427	-1.31	-4.368	-1.09	-4.436	-1.23

Table 3. *Similarity scores as a function of Veracity in the small talk condition*

Similarity scores	Truth Tellers		Liars		<i>F</i>	<i>p</i>	<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>			
Hands and fingers movements	57.94	27.68	70.58	27.33	2.23	.14	-.46
One arm movements	41.52	32.16	30.33	24.57	.35	.55	.39
Two arms movements	59.77	37.98	57.58	35.78	.00	.96	.06
Spatial details	14.69	30.79	33.66	42.93	.09	.76	-.50
Temporal details	40.98	30.57	54.60	30.23	1.518	.23	-.45
Visual details	10.36	12.88	13.20	23.68	.21	.65	-.14
Audio details	41.17	50.73	33.93	45.58	.42	.52	.15
Action details	45.50	22.21	55.45	27.87	.78	.39	-.39