

## Appendix

For further exploratory analyses, we added response time, the number of deletions, and gaps between key-press events as dependent variables

### *Para-linguistic keyboard-related variables*

Following Zhou's (2005) taxonomy of deceptive behavior, we measured the para-linguistic keyboard-related variables for participants' answers to the ten main questions. The response time in milliseconds from the first key-press event to finishing the question, the number of occasions the backspace/delete key was pressed, the number of gaps between key-press events that were longer than 100ms, 200ms, and 300ms. These variables were intended to represent participants' editing behavior and have previously been shown to be informative (Derrick et al., 2013; Zhou, 2005).

Specifically, we recorded the following para-linguistic variables: response time in milliseconds, number of deletions, number of gaps between key-presses larger than 100ms, 200ms and 300ms. All of these three non-verbal measures might be a way to model cognitive processes such as cognitive load involved in formulating answers through non-verbal behavior in a non-interactive setting (e.g., Derrick et al., 2013; Zhou, 2005). We conducted separate one-way ANOVAs with Veracity as factor. There were no significant effects of Veracity on the average response time ( $M_{truthful} = 2953$ ,  $SD_{truthful} = 2595$ ;  $M_{deceptive} = 2984$ ,  $SD_{deceptive} = 2606$ ); on the average number of deletions ( $M_{truthful} = 12.57$ ,  $SD_{truthful} = 11.74$ ;  $M_{deceptive} = 13.16$ ,  $SD_{deceptive} = 13.72$ ); nor on the average number of key-press gaps larger than 100ms ( $M_{truthful} = 120.91$ ,  $SD_{truthful} = 47.45$ ;  $M_{deceptive} = 128.81$ ,  $SD_{deceptive} = 58.91$ ). However, for the average number of key-press gaps larger than 200ms ( $M_{truthful} = 62.79$ ,  $SD_{truthful} = 30.65$ ;  $M_{deceptive} = 69.84$ ,  $SD_{deceptive} = 33.03$ ) and larger than 300ms ( $M_{truthful} = 35.10$ ,  $SD_{truthful} = 19.90$ ;  $M_{deceptive} = 40.04$ ,  $SD_{deceptive} = 22.74$ ), there were significant differences,  $F(1, 352) = 4.30$ ,  $p = .038$ ,  $f = 0.11$ , and  $F(1, 352) = 4.68$ ,  $p = .031$ ,  $f = 0.11$ , respectively. Although these differences might be an indicator of participants' hesitation when formulating answers and might then be in line with the cognitive load rationale, these conclusions are merely tentative and merit further replication in future studies. Furthermore, the effect sizes are small and thereby as single cues to deception, these para-linguistic variables are of limited relevance for the detection of individual cases.