Moritz Maleck and Tom Gross*

AnswerTruthDetector: a combined cognitive load approach for separating truthful from deceptive answers in computer-administered questionnaires

Abstract: In human-computer interaction, much empirical research exists. Online questionnaires increasingly play an important role. Here the quality of the results depend strongly on the quality of the given answers, and it is essential to distinguish truthful from deceptive answers. There exist elegant single modalities for deception detection in the literature, such as mouse tracking and eye tracking (in this paper, respectively, measuring the pupil diameter). Yet, no combination of these two modalities is available. This paper presents a combined approach of two cognitive-load-based lie detection approaches. We address study administrators who conduct questionnaires in the HCI, wanting to improve the validity of questionnaires.

Keywords: truth detection; lie detection; questionnaire validation; eye tracking; mouse movements; cognitive-load-based deception detection

1 Introduction

Deceptive answers in questionnaires in human-computer interaction (HCI) decrease the quality of study results. Deception in questionnaires can, among other things, be a result of social desirability, which means that participants explicitly answer questionnaires in such a way that their answers are perceived as positively as possible by others and do not correspond to their true convictions [1]. Honesty declarations (e.g., as part of a briefing, or an informed consent) can be used to reduce dishonesty [2]; yet, such declarations have not always shown effect at all [3].

Many authors in HCI and other fields have dealt with deceptive answers in questionnaires as well as their detection [4–11]. Krumpal [7] have generally analysed various reasons for misreporting in surveys and came to the conclusion that the design of a survey can lead to more honest answers and therefore higher reliability of the data. Preisendörfer and Wolter [6] have compared the effect of dishonesty in face-to-face surveys compared to mail surveys. While mail surveys tended to be answered more truthfully, an interesting side finding was that truthfully answered surveys were answered with less delay than deceptive ones.

More specifically, related work has suggested approaches to accurately distinguish truthful from deceptive answers from participants. Fang, Sun, Zheng, Wang, Deng and Wang [4] have successfully applied eye tracking to detect deceptive answers in questionnaire surveys to increase the reliability of study data. Mazza, Monaro, Burla, Colasanti, Orrù, Ferracuti and Roma [10] have applied mouse tracking and machine learning to detect deceptive answers. Eye tracking and mouse tracking techniques evolved from invasive approaches to cognitive-load-based ones. Cognitive-based lie detection techniques form a valid alternative to invasive stress-based deception detection. Invasive stress-based methods, as traditional approaches, try to detect lies by measuring "changes in blood pressure, heart rate and respiration rate" [12] using a polygraph. An increasing field is the research in cognitive-load-based approaches. Detecting deception based on the paradigm of lies being accompanied by higher cognitive effort [13] has successfully been applied by the use of eye tracking as well as mouse tracking: Pupil diameters in eye tracking data [14, 15] as well as mouse movements [16, 17] are valid indicators for measuring the real-time cognitive load of individuals. Furthermore, both pupil diameters [4, 18] and mouse movements [19, 20] measures have already been successfully used separately to reveal deceptive answers in online questionnaires. Most recently, mouse tracking-based deception research has been supported by machine learning [10, 21].

---

*Corresponding author: Tom Gross, Human-Computer Interaction Group, University of Bamberg, Bamberg, Germany,
E-mail: hci@uni-bamberg.de
Moritz Maleck, Human-Computer Interaction Group, University of Bamberg, Bamberg, Germany

Relying on eye tracking and mouse tracking per se – that is, a single modality – comes with various limitations. For instance, in eye tracking, uncontrollable person-dependent factors (like iris brightness, skin brightness, pupil size, eyelashes or eyelids, and drying of the eyes during a study) have been challenging [22]. With mouse tracking, there is also an age-group-related differentiated usage performance [23], which can lead to varying quality of the results. Moreover, the use of a single modality is not appropriate in terms of accessibility, as a person-related valid application of both modalities cannot always be guaranteed.

In this paper, we present and demonstrate a novel approach for separating truthful from deceptive answers in computer-administered questionnaires that combines two cognitive load-based lie detection techniques for computer-administered questionnaires: mouse tracking and eye tracking. Figure 1 gives an impression of the AnswerTruthDetector Questionnaire setup and the AnswerTruthDetector Detector Tool.

In particular, this paper has three contributions:

1. A combination of various eye tracking and mouse tracking features for separating truthful from deceptive answers in computer-administered questionnaires.
2. A continuous universal truth score for eye tracking, mouse tracking, and the combined approach allowing a more fine-granular prediction of the truthfulness of an answer compared to a binary-only classification (please note that in this paper, we are interested in identifying reliable answers in empirical data and so we take a positive perspective and want to identify truthful answers).
3. An extensive toolbox for conducting computer-administered questionnaires and separating truthful from deceptive answers in these.

2 Related work

Invasive stress-based detection methods like traditional lie detection techniques rely on body sensors that measure the three main channels "cardiovascular activity, respiratory activity and electrodermal activity" [12]. Fear and stress (optimally caused by a lie) can be observed by using a polygraph. Research and practice have mainly focused on this technique for over a century [12, 24]. However, polygraph truth classifications have often been criticised [25].

Cognitive-load-based deception detection methods are an alternative to stress-based methods. They underlie a basic yet simple principle: When lying, the cognitive effort is higher than when telling the truth due to a broad range of reasons [13]. There are several possibilities to measure an individual's cognitive load, yet all non-invasive and easier implementable in the HCI context in contrast to stress-based lie detection methods that require sensors attached to the body. Most commonly, cognitive dynamics and decision processes can be measured using the response time [26], eye tracking [14, 27], and mouse tracking [16, 17, 28]. In contrast to eye tracking and mouse tracking, response time does not allow zooming into the different stages of a process [29].

Eye tracking exists of various features that have already been applied in measuring cognitive processes and detecting deceptive behaviour, whereby the pupil diameter is best suited for detection compared to other features. They give a detailed real-time insight into mental processes and represent a robust and reliable data source since it is not possible to influence the size of the pupils consciously and voluntarily [14]. Another disadvantage of eye tracking, in general, is the unequal cross-user applicability, which we provide a solution for with the combined approach. Traditional eye tracking looks at eye movements, including
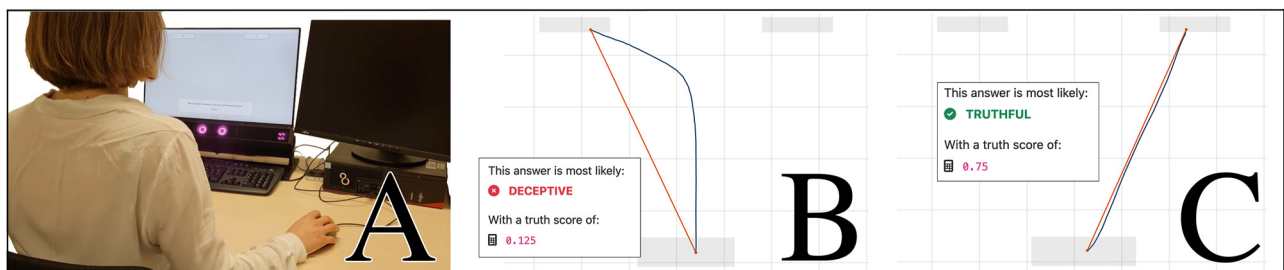

Figure 1: AnswerTruthDetector in action. (A) Setup with a study participant answering questions in the AnswerTruthDetector questionnaire tool. (B) AnswerTruthDetector detector tool screenshot of mouse trajectory and truth score of deceptive answer. (C) AnswerTruthDetector detector tool screenshot of mouse trajectory and truth score of truthful answer.

saccades and fixations. Research exists to use fixations as indicators for increased cognitive load [30, 31] as well as for deceptive answers [4, 32]. Yet, a reliable and stable indicator to measure cognitive load is the pupil diameter. The pupil diameter does not only change because of light changes but also reflects mental processes and workload [14, 31, 33]. An increase in diameter can be an indicator for, among other things, emotional reactions, mental (cognitive) load and decision processes [14]. It has been found that the pupil of a participant dilates before, while and after performing a deceptive action [34]. The effect of greater pupil dilations when lying could be observed in various studies [24, 35–37]. This has also been applied in computer-administered questionnaires [4, 18]. Another finding is that the pupil diameter can predict upcoming yes/no decisions even before answering, with greater pupils for forthcoming yes answers than for no answers [38].

Mouse tracking is used to evaluate answers regarding their truthfulness. Like eye tracking, mouse tracking gives a detailed view into the real-time mental processes of an individual. Mouse movements can therefore be used as valid indicators to detect higher cognitive load [16, 17, 39] and competing stimuli to the brain [19]. A differentiation between truthful and deceptive answers by analysing mouse movements is possible [40]. To achieve that, the mouse trajectory for a specific question can be analysed: The straighter the travelled path from the start point to a selected answer, the less distractive the opposite answer was to the participant [20]. To maximise the distance of the mouse to be travelled, most commonly, the participant is required to start moving the mouse at the bottom of the screen and move the cursor to the answer options that are placed at the top opposite corners [28]. Travelled trajectories can be analysed spatially and temporally. Widely used are the area under the curve (AUT) and maximum deviation (MD); with the AUT being the geometrical area between the actual trajectory of a user and the direct, straight line connecting the start- and end-point; and the MD being the maximum deviation between these two paths [28]. Another possibility to measure cognitive dynamics is counting the x-flips of a trajectory – that is, the direction changes [41]. A novel approach is the maximum log ratio (MLR), which is the maximum "ratio of the target distance to the alternative distance" [42] of alternative answer options. For all those features, higher values show a higher deviation towards the alternative, not selected response [42] and, so, a higher probability for a deceptive answer. In the context of detecting untruthful answers in online questionnaires, current research often focuses on relying on machine learning algorithms to learn typical mouse movement behaviour of users to classify the truthfulness of a question based on the derived model [10, 21, 43]. Another, more mass deployable concept uses a pre-defined model as to separate truthful from deceptive information [20]. A general disadvantage of applying mouse tracking for detection is that participants are required to move their mouse unintuitively due to the arrangement of the window frames and rules (e.g., to move the mouse immediately after the question presentation). Another limitation of mouse tracking is the applicability to binary answer types (especially yes/no question types) only. In addition, there can be differences in quality across different persons [23].

In this article, we aim to build a solid fundament to increase the quality and validity of answers in online questionnaires to support study administrators. Eye and mouse tracking as single modalities already offer great possibilities to detect deceptive answers in pre-study questionnaires and therefore, to reformulate such sensible parts in final studies. Yet, both modalities have their limitations and may under certain circumstances fail to work properly. Instead of only relying on one single modality, we therefore propose to combine both – so that one modality can support the other in case of temporal failures. To the best of our knowledge, elegant combinations of mouse tracking and eye tracking for truth and deception differentiation are missing, which can be explained by the heterogeneity of requirements of both modalities that are considered. A combination is particularly suitable as both modalities are based on a similar principle. For instance, mouse tracking can be seen as a cheaper but not less reliable alternative to eye tracking [44] and therefore complements and supports particularly well the eye tracking modality in case of unreliability, and vice versa. The use of both modalities can also be an indicator of the plausibility of a single modality. Gross deviations of the respective other modality may indicate incorrect classifications. Existing combinations of other modalities have already demonstrated significantly higher reliability and efficiency compared to the use of a single modality [45, 46]. Combining modalities is also a step towards time-saving automation of distinguishing truthful from deceptive responses [46], as less extensive analysis of individual results by study administrators is required. That is, because incorrect classifications can be revealed more easily by comparing them with other modalities used, and thus automated correction steps can be applied.

3 A combined approach for truth detection

Mouse-tracking and eye-tracking modalities are a solid basis for differentiating truthful from deceptive answers in computer-administered questionnaires. To combine modalities and their features, we introduce a uniform per-question and -participant truth score. This score measures each modality allowing their combination. Study administrators can configure the combination through optimal weighting of the individual modalities and respective features. Based on pre-defined models as well as the analysis of participant-specific behaviour, our approach is fast and easily deployable in contrast to machine learning approaches, which need expensive and time-consuming training of a model. To combine eye tracking and mouse tracking, we leverage the particular strengths of both approaches.

We apply mouse tracking to reveal an answer's truthfulness by considering different requirements for gathering data and combining various detection features for separating truthful from deceptive mouse trajectories. Mouse tracking requires a specific design of the questionnaire tools. This way, we use mouse trajectories to classify answers as truthful or deceptive. Furthermore, various detection features are available, we consider four spatial features (area under the trajectory curve, maximum deviation, maximum log ratio, and x-coordinate flips). Figure 2 visualises three previously named features (AUT, MD, MLR). For all features, we provide a binary truth deception classification (a) based on fixed thresholds, and (b) based on the default behaviour of a participant. In addition, we apply configurable, fixed thresholds (e.g., trajectories with more than one x-coordinate flip are considered deceptive and vice versa). This way, our approach allows the application of pre-defined models. Furthermore, we consider the default behaviour of a participant (e.g., the participant has

a mean of three x-coordinate flips because of trembling mouse movements; therefore, the fixed threshold of one x-coordinate flip would make less sense). Thus, for all mouse tracking features, (a) and (b) are available and weightable within our combined approach. Furthermore, we generate a binary truth classification for each feature and finally result in a combined classification (truth score).

We consider eye tracking for separating truthful from deceptive answers by looking at the participants' continuously measured pupil diameters as a window into mental processes. Pupil diameters reveal significant insights into real-time mental processes. Based on the paradigm that lying requires a higher cognitive load than truth-telling, we consider pupils dilating more when giving deceptive answers than truthful answers. By applying this knowledge, we compare the pupil size for a specific question to a participant's default (mean) pupil diameter. If the diameter is higher than average for a question, we detect this question to have caused a higher cognitive effort than other questions. Therefore, we consider the question as a lied item. We further introduce a different, novel approach for lie detection by applying the finding that pupil diameters predict upcoming yes/no answers. We detect probably deceptive answers by comparing the expected yes/no response with the finally selected answer of a participant; in case of a contradiction, we assume a higher probability for a deceptive answer. We provide a binary truth-deception classification for each eye-tracking feature (higher cognitive load detection, answer prediction). As already for mouse tracking, we further calculate a combined classification.

We introduce a continuous universal truth score for combining the two modalities and, six different features. It allows a comparison of the individual features as well as the modalities. A continuous score between 0 and 1 allows more fine granular classifications (i.e., higher values shall indicate a higher probability for an answer to have been
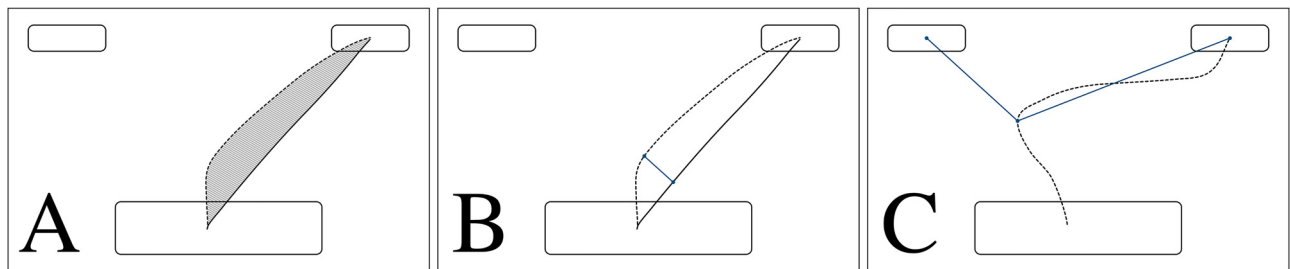

Figure 2: AnswerTruthDetector detector tool's three spatial mouse trajectory visualisations when answering a question. The participant starts moving the mouse at the bottom centre of the screen to the answer options placed in the upper corners. (A) Area under the trajectory curve, (B) maximum deviation, (C) maximum log ratio.

answered truthfully and vice versa). With such a universal truth score for both modalities available, we provide an easily applicable combination. The truth score is available for each feature and modality, but most importantly, as a combined truth score of all features for a specific question. We further provide the possibility for weightings that can be applied for each modality to enable different weights for different study participants, as we have seen previously that, in particular, eye tracking does not work equally well for every individual. This way, we reach a higher diversity by decreasing the possible number of participants to be excluded from a study. This would be impossible for a binary truth classification of a single modality.

We provide an overview of the logic behind our concept in Figure 3. Based on the combined approach truth score, the combined approach makes it possible to easily obtain the probability for a specific question to have been answered truthfully or deceptive.

4 AnswerTruthDetector: questionnaire tool and detector tool

In this section, we present the implementation and functionality of the AnswerTruthDetector. The AnswerTruthDetector is designed as a system that includes the Questionnaire Tool and the Detector Tool as central components (cf. Figure 4). The Questionnaire Tool runs under Python
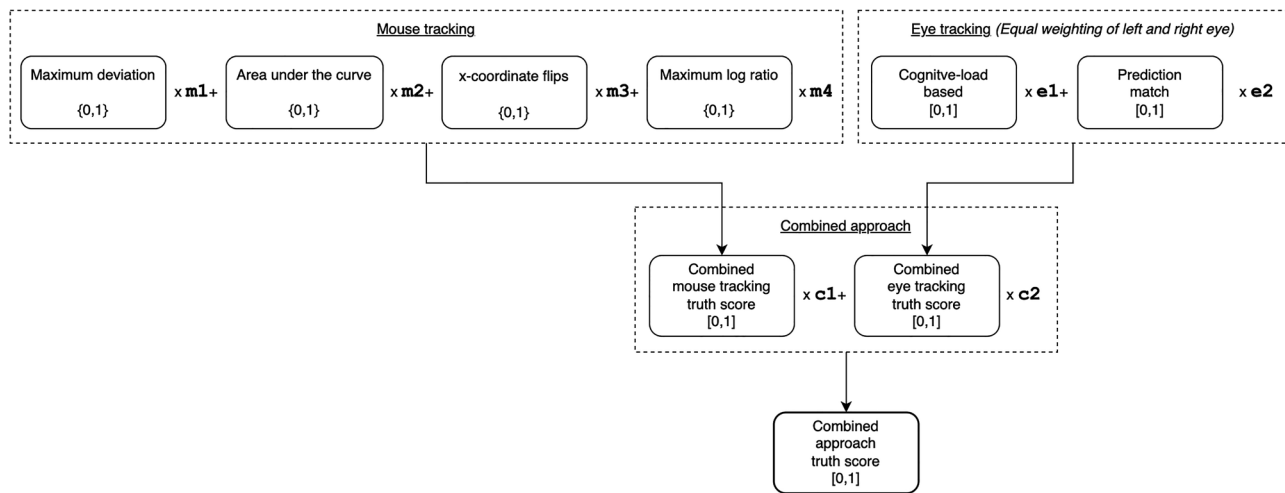

Figure 3: Both modalities (mouse tracking and eye tracking) with their six functions (four mouse tracking functions, two eye tracking functions) result in the combined truth score, which indicates the probability of answering a question truthfully. *m1-4*, *e1-2* and *c1-c2* allow the weighting of the different functions and modalities. To determine these, study results of the respective participant are considered (e.g., if the quality of the eye tracking data is good, a higher weighting of this modality can be chosen than if it is poor).
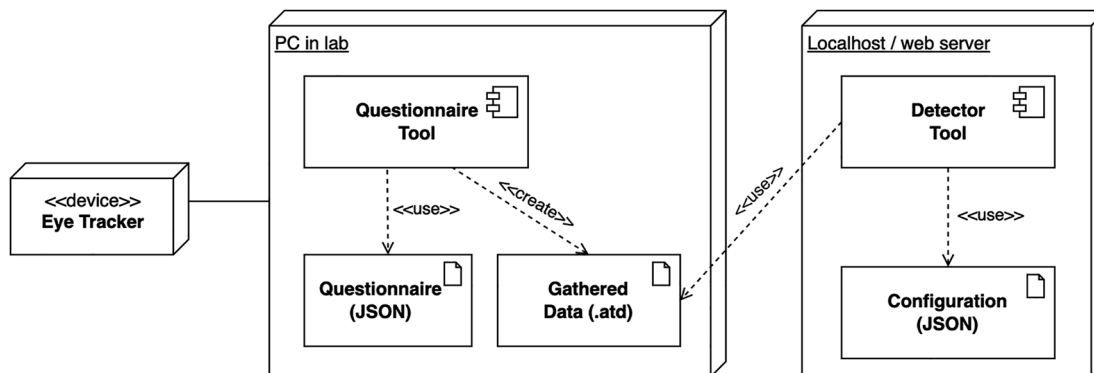

Figure 4: Overview of the AnswerTruthDetector system with its two components – the questionnaire tool and the detector tool. The questionnaire tool imports the questionnaire as a JSON file and provides an export of the gathered data (selected answers, mouse tracking data, eye tracking data) in a program-specific file format. This file is imported within the detector tool and is then used for the analysis based on an imported configuration file.

3.8.10, Windows 10 or higher and with a connected Tobii Pro Spectrum eye tracker (note that only this specific prototype is implemented with a Tobii eye tracker, but can easily be adapted for usage of other eye trackers). The Detector Tool is a universal web application based on the JavaScript Svelte Framework (version 3.48.0).

4.1 Gathering data with the questionnaire tool

The Questionnaire Tool is optimally designed for both mouse tracking and eye tracking modalities, both requiring specific design implications. By considering and merging different requirements of each eye tracking and mouse tracking modality into one tool, the AnswerTruthDetector Questionnaire Tool allows a universal application.

To meet the requirements for mouse tracking, a special arrangement of the questionnaire components (question and answer options) is used to achieve a maximum distance for the mouse trajectories. This includes forcing the participant to start moving the cursor at the bottom of the screen, which is done by requiring the participant to click a button placed at the bottom centre of the screen to show the following question (cf. Figure 5). Furthermore, to prevent the participant from moving the mouse only after completing the thinking process, the Questionnaire Tool shows an alert after selecting the answer, asking to start moving the mouse faster if the cursor was not moved within a certain amount of time at the very beginning.

Two functionalities are implemented to achieve the best possible quality of the logged eye-tracking data. Firstly, before each question presentation, the participant must look at a fixation cross to ensure valid measurements for a specific question right from the beginning. Secondly, each time between selecting an answer option and presenting the next question, a fixed (configurable) time delay is applied so that the pupil has time to return to its baseline size again.

The study administrator of a questionnaire can import the questions to be used for the questionnaire as a JSON file which is standardised for the AnswerTruthDetector toolbox. For research reasons, for each question, it can be further stated whether a participant should lie or not, whereby the fixation cross would be replaced by a letter (T = Truth, L = Lie). In addition, the order of the questions and the optional lie information can be randomised.

The Questionnaire Tool applies eye tracking and mouse tracking. By default, it logs eye coordinates and pupil diameters for each left and right eye. For each log entry, the tool stores where the participant looked at (question, answer option, screen). Furthermore, it logs the mouse coordinates of a participant. To enable further usage and evaluation of these data, the tool stores all coordinates of the window and the displayed frames for each question.

After the participant has finished the questionnaire, the study administrator can export the logged data in a program-specific file format to then import it for further evaluation in the Detector Tool. Again, this is possible without delay so that a quasi-real-time truth classification of given answers is possible.

4.2 Separating truthful from deceptive answers with the detector tool

The second tool of the AnswerTruthDetector toolbox is the Detector Tool. The Detector Tool allows the study administrator to separate truthful from deceptive answers in questionnaires conducted using the Questionnaire Tool from the previous section.

The truth score for each question is calculated based on the features and their corresponding logics presented above (for mouse tracking: area under the trajectory curve, maximum deviation, x-coordinate flips, maximum log ratio; for eye tracking: pupil diameter to measure cognitive load, pupil diameter to predict yes/no answers).
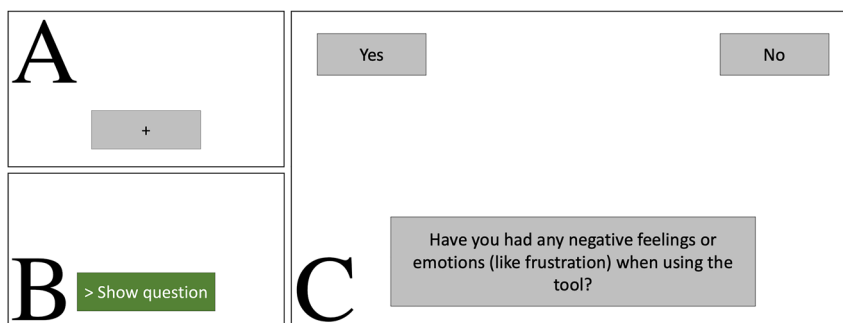

Figure 5: Scheme of AnswerTruthDetector questionnaire tool's GUI showing the typical sequence for each question: (A) the participant looks at a fixation cross for a pre-defined time, (B) the participant then clicks on show question, (C) then the question and the answer options are displayed.

Mouse tracking detection can rely on pre-defined and/or per-participant behaviour. The study administrator can apply a pre-defined model. The tool can rely on the default behaviour of a participant – that is, it can be considered if the typical mouse trajectory of a participant differs from the defined model. Finally, the study administrator can set unique weightings (defined model vs. participant behaviour) per participant. Applying a pre-defined model would not make sense for the two eye tracking features because each pupil is unique and cannot be compared over multiple participants. Therefore, for both features, the default behaviour of the participant during the questionnaire is used.

For eye tracking and mouse tracking, we calculate a continuous truth score based on the classifications and weightings of each single feature. The combined truth score brings both the two modalities and their features together considering the per-participant-weightings. All weightings (plus other configuration options) can be changed in a configuration file imported by the Detector Tool.

Based on the combined truth score, the study administrator can easily obtain the probability that a selected question has been answered truthfully or deceptively (cf. Figure 1B and C). Owed to the central function of this component, this is also the start view after selecting a participant and question. For eye tracking and mouse tracking modalities, more detailed information is available to the study administrator to understand the classifications better. For the eye tracking modality (cf. Figure 6), the study administrator can see the binary truth classification for each feature and each left and right eye. For the answer-prediction feature, the study administrator can also see which answer was expected and selected. Furthermore, all mean pupil diameters (current question, all questions, while reading the questions, baseline pupil size) are displayed, and the pupil diameter curve for both the left and right eye is available. The resulting truth score for all eye tracking features is also printed.

In the mouse tracking section for a selected question (cf. Figure 7), the tool lists all four features (including results for the current question, mean values over the total questionnaire and a per-feature truth score), as well as the resulting truth score for mouse tracking. In addition, for a better understanding of the values, the drawn mouse trajectory curve of the participant is available in an interactive diagram, where the different features can be visualised.

All data of the AnswerTruthDetector Detector Tool can also be exported to the Excel format.
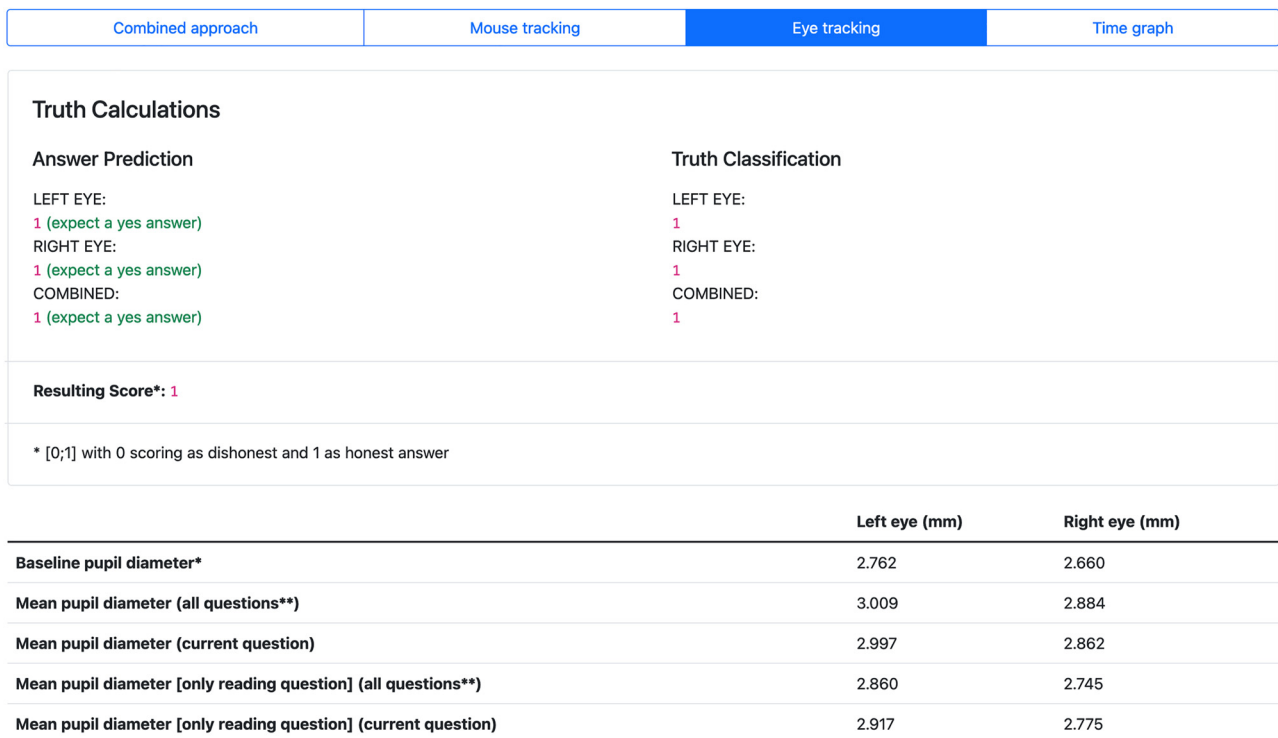

Combined approach	Mouse tracking	Eye tracking	Time graph

Truth Calculations

Answer Prediction

LEFT EYE:
1 (expect a yes answer)
RIGHT EYE:
1 (expect a yes answer)
COMBINED:
1 (expect a yes answer)

Truth Classification

LEFT EYE:
1
RIGHT EYE:
1
COMBINED:
1

Resulting Score*: 1

* [0;1] with 0 scoring as dishonest and 1 as honest answer

	Left eye (mm)	Right eye (mm)
Baseline pupil diameter*	2.762	2.660
Mean pupil diameter (all questions**)	3.009	2.884
Mean pupil diameter (current question)	2.997	2.862
Mean pupil diameter [only reading question] (all questions**)	2.860	2.745
Mean pupil diameter [only reading question] (current question)	2.917	2.775

Figure 6: Screenshot of AnswerTruthDetector detector tool depicting the eye tracking data analysis.

Combined approach	Mouse tracking	Eye tracking	Time graph

Truth Calculations

Mouse-Tracking Measures	Result for current question	Mean (all questions)*	Median (all questions)*	Truth Score**
Area Under the Trajectory Curve	12955.491	99244.668	–	1 ?
Maximal Deviation	48.423	301.702	–	1 ?
Maximal Log Ratio	0.945	7.992	–	1 ?
X-coordinate Flips	0	1.700	1	1 ?
Resulting truth score				1 ?

* without the first 0 questions (= calibration questions)
** [0;1] with 0 scoring as dishonest and 1 as honest answer

Figure 7: Screenshot of AnswerTruthDetector detector tool depicting the mouse tracking data analysis.

4.3 Possible applications

Our AnswerTruthDetector supports study administrators in conducting studies. We focus on a preventive application – that means we intend the AnswerTruthDetector to be used in preliminary studies to identify the need for optimisation in future main studies. The tool is therefore not meant to be used for filtering out answers in final studies, since this could result in a possible distortion of the results.

We propose a (non-exhaustive) list of possible usage scenarios of the AnswerTruthDetector:

(a) The AnswerTruthDetector can be used in a lab. Participants fill out questionnaires on a computer with at least a conventional mouse, and a (semi-)professional eye tracker connected. Depending on the eye-tracker used, their head can be free to move. The quality of the results depends on the concrete eye tracker used. Collected data can be analysed and be used for further questionnaire-optimisation iterations.

(b) The AnswerTruthDetector can be used on a conventional laptop or PC outside a lab. Minimum requirements are a mouse or a trackpad, and a high-quality (built-in or external) face-camera. The quality of the data with a conventional face-camera is worse than with an eye tracker, still it provides helpful insights for studies without an available eye tracker and/or low budget. For using a face-camera as eye tracking device, it might be necessary to install additional software. Participants shall be instructed to answer the questionnaire in a quiet room without any disturbing distractions. The collected data can then also be used to further optimise questionnaires.

The AnswerTruthDetector helps to reduce the number of sensitive questions with a high probability of being answered deceptively. For that, the AnswerTruthDetector can be used to conduct preliminary studies to identify problematic questions by involving a small number of test participants. These findings can then be used to optimise the questionnaire for the main study and in this way achieve a better quality of the study results.

Two technical implementations of the AnswerTruthDetector are possible: system (*iS*) versus plugin (*iP*). Apart from using the AnswerTruthDetector system with its two tools (*as is*), it is further possible to provide it for other platforms for seamless integration as a plugin. This allows an implementation of the logic of the Detector Tool as an R library or Python library in existing system landscapes. Studies and questionnaires can be conducted this way with existing questionnaire applications (under the condition that they gather eye and mouse tracking data in sufficient quality). In data analysis applications such as SPSS, the core functions of the Detector Tool can then be used as a plugin solution.

5 Conclusions and future work

We introduced a novel approach for combining eye tracking and mouse tracking modalities to separate truthful from deceptive answers in computer-administered questionnaires in the context of cognitive-load-based deception detection. Our universal truth score provides a fine-granular prediction of the truth content of an answer compared to binary-only classifications. We implemented six

different features (two eye tracking features and four mouse tracking features) for the combined approach. We reach a maximum diversity by allowing extensive per-participant and per-study weightings of all features and modalities. This way, possible differences among the participants can be considered. We implemented and demonstrated the approach within the AnswerTruthDetector toolbox, providing a Questionnaire Tool and a Detector Tool, both offering a broad range of functionalities to the study administrator.

However, there are also limitations that need to be considered when using the two modalities mouse tracking and eye tracking. For both modalities it is important to keep in mind that there can be significant differences in the quality of the data depending on the individual participant, although we address this with our introduced combination. Furthermore, it must be considered that with eye tracking the pupil size is dependent on lighting conditions. Therefore, it is necessary to carry out the test in a laboratory to reduce unwanted effects as much as possible. As we explicitly distinguish truthful from deceptive answers, we must take into account the possibility of different reasons for higher cognitive effort on a question (e.g., conscious vs. unconscious lying, not understanding a question, question too difficult to answer, language barriers). In this regard, the pupil diameter, which is particularly used in our concept, can not only increase due to deceptive behaviour but also due to other factors such as mental effort, emotion, uncertainty, or urgency. A more fine-grained classification of reasons will be addressed in the future.

Existing combinations of other modalities from related work have shown to bring improvements compared to the usage of a single modality. Since eye and mouse tracking as single modalities already had good performance in detecting deceptive answers, we build upon key findings from other combinations and adapt these to our concrete modalities. This way, study administrators will have to do less time-consuming analysis of the data, since incorrect classifications can be detected more easily and beyond that, can even be corrected automatically. Combining modalities per se is promising. Compared to related work, we combine two modalities which have not been combined until now; whereby these two modalities can be used very easily (possible even with only a conventional mouse and face-camera). Moreover, the applied modalities do not disturb the participants since they can behave as they normally would (without having to wear sensors, keeping the head fixated, etc.). The concept can be applied without further training of a data model, since by default the measured typical behaviour of a participant is used as basis to detect outliers. Configuration options, such as the weighting of the concrete

modalities, make the concept highly adaptable. Since in our tests of our approach we told the study participants when to give true and when to give deceptive answers, we got a clear understanding that our approach indeed measures deception. Therefore, the internal validity can be considered high [47]. Also the approach can be applied for various types of online questionnaires in the sense of high external validity. Our tests reliably showed similar results across participants and study setups.

We have tested the AnswerTruthDetector with 14 (5 female, 9 male, 0 diverse) participants. Their age ranged from 19 to 40 years old. The participants were recruited with mouth-to-mouth sampling. Participants were asked to answer 16 questions using the AnswerTruthDetector Questionnaire Tool in a lab. After the informed consent, they were instructed to answer some questions truthfully and some deceptively. We were overall satisfied with the evaluation results of the AnswerTruthDetector Detector Tool. Yet, we found a correlation between the number of deceptive answers of individual participants and the classification quality – the classification quality was higher for participants who had more truthful answers.

In the future, we plan a systematic user study of the AnswerTruthDetector with a new sampling rate of 1200 Hz (originally it was 60 Hz). Future versions of the system could provide adequate suggestions for weightings. We also plan to introduce mechanisms to identify different reasons for deceptive answers.

References

1. Olson J. S. *Ways of Knowing in HCI*; Springer New York: New York, NY, 2014; p. 1 Online-Ressource (XI, 472 p.): Ill.

2. Shu L. L., Mazar N., Gino F., Ariely D., Bazerman M. H. Signing at the beginning makes ethics salient and decreases dishonest self-reports in comparison to signing at the end. *Proc. Natl. Acad. Sci. U. S. A.* 2012, *109*, 15197−15200.

3. Isoni A., Read D., Kolodko J., Arango-Ochoa J., Chua J., Tiku S., Kariza A. Chapter 4.4 − can upfront declarations of honesty improve anonymous self-reports of sensitive information? In

Dishonesty in Behavioral Economics; Bucciol A., Montinari N., Eds.; Academic Press: London, 2019; pp. 319−340.

4. Fang X., Sun Y., Zheng X., Wang X., Deng X., Wang M. Assessing deception in questionnaire surveys with eye-tracking. *Front. Psychol.* 2021, *12*, 774961.

5. Hyman H. Do they tell the truth? *Publ. Opin. Q.* 1944, *8*, 557−559.

6. Preisendörfer P., Wolter F. Who is telling the truth? A validation study on determinants of response behavior in surveys. *Publ. Opin. Q.* 2014, *78*, 126−146.

7. Krumpal I. Determinants of social desirability bias in sensitive surveys: a literature review. *Qual. Quantity* 2013, *47*, 2025−2047.

8. Nederhof A. J. Methods of coping with social desirability bias: a review. *Eur. J. Soc. Psychol.* 1985, *15*, 263−280.

9. Clark S. J., Desharnais R. A. Honest answers to embarrassing questions: detecting cheating in the randomized response model. *Psychol. Methods* 1998, *3*, 160−168.

10. Mazza C., Monaro M., Burla F., Colasanti M., Orrù G., Ferracuti S., Roma P. Use of mouse-tracking software to detect faking-good behavior on personality questionnaires: an explorative study. *Sci. Rep.* 2020, *10*, 4835.

11. Bond C. F., DePaulo B. M. Accuracy of deception judgments. *Pers. Soc. Psychol. Rev.* 2006, *10*, 214−234.

12. Synnott J., Dietzel D., Ioannou M. A review of the polygraph: history, methodology and current status. *Crime Psychol. Rev.* 2015, *1*, 59−83.

13. Walczyk J. J., Harris L. L., Duck T. K., Mulay D. A social-cognitive framework for understanding serious lies: activation-decision-construction-action theory. *New Ideas Psychol.* 2014, *34*, 22−36.

14. Eckstein M. K., Guerra-Carrillo B., Miller Singley A. T., Bunge S. A. Beyond eye gaze: what else can eyetracking reveal about cognition and cognitive development? *Dev. Cogn. Neurosci.* 2017, *25*, 69−91.

15. Sirois S., Brisson J. Pupillometry. *WIREs Cogn. Sci.* 2014, *5*, 679−692.

16. Freeman J., Dale R., Farmer T. Hand in motion reveals mind in motion. *Front. Psychol.* 2011, *2*, 59.

17. Rheem H., Verma V., Becker D. V. Use of mouse-tracking method to measure cognitive load. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* 2018, *62*, 1982−1986.

18. Webb A. K., Hacker D. J., Osher D., Cook A. E., Woltz D. J., Kristjansson S., Kircher J. C. *Eye Movements and Pupil Size Reveal Deception in Computer Administered Questionnaires*; Schmorrow D. D., Estabrooke I. V., Grootjen M., Eds.; Springer: Berlin, 2009; pp. 553−562.

19. Duran N. D., Dale R., McNamara D. S. The action dynamics of overcoming the truth. *Psychonomic Bull. Rev.* 2010, *17*, 486−491.

20. Jenkins J., Proudfoot J., Valacich J., Grimes G., Jay F., Nunamaker J. Sleight of hand: identifying concealed information by monitoring mouse-cursor movements. *J. Assoc. Inf. Syst.* 2019, *20*, 1−32.

21. Monaro M., Fugazza F. I., Gamberini L., Sartori G. *How Human-Mouse Interaction can Accurately Detect Faked Responses About Identity*; Springer International Publishing: Berlin, 2017; pp. 115−124.

22. Schnipke S. K., Todd M. W. Trials and tribulations of using an eye-tracking system. In *CHI'00 Extended Abstracts on Human Factors in Computing Systems*: The Hague, The Netherlands, 2000; pp. 273−274.

23. Hertzum M., Hornbæk K. How age affects pointing with mouse and touchpad: a comparison of young, adult, and elderly users. *Int. J. Hum. Comput. Interact.* 2010, *26*, 703−734.

24. Dionisio D. P., Granholm E., Hillix W. A., Perrine W. F. Differentiation of deception using pupillary responses as an index of cognitive processing. *Psychophysiology* 2001, *38*, 205−211.

25. Kleinmuntz B., Szucko J. J. On the fallibility of lie detection. *Law Soc. Rev.* 1982, *17*, 85−104.

26. Bergert F. B., Nosofsky R. M. A response-time approach to comparing generalized rational and take-the-best models of decision making. *J. Exp. Psychol. Learn. Mem. Cognit.* 2007, *33*, 107−129.

27. Zénon A. Eye pupil signals information gain. *Proc. R. Soc. Biol. Sci.* 2019, *286*, 20191593.

28. Freeman J. B., Ambady N. MouseTracker: software for studying real-time mental processing using a computer mouse-tracking method. *Behav. Res. Methods* 2010, *42*, 226−241.

29. Koop G. J., Johnson J. G. Response dynamics: a new window on the decision process. *Judgm. Decis. Mak.* 2011, *6*, 750−758.

30. Zagermann J., Pfeil U., Reiterer H. *Measuring Cognitive Load Using Eye Tracking Technology in Visual Computing*; ACM Press: N.Y., 2016.

31. Zagermann J., Pfeil U., Reiterer H. Studying eye movements as a basis for measuring cognitive load. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*: Montreal, QC, Canada, 2018; pp. 1−6.

32. Cook A. E., Hacker D. J., Webb A. K., Osher D., Kristjansson S. D., Woltz D. J., Kircher J. C. Lyin' eyes: ocular-motor measures of reading reveal deception. *J. Exp. Psychol. Appl.* 2012, *18*, 301−313.

33. Pfleging B., Fekety D. K., Schmidt A., Kun A. L. A model relating pupil diameter to mental workload and lighting conditions. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*: San Jose, California, USA, 2016; pp. 5776−5788.

34. Wang J. T.-Y., Spezio M., Camerer C. F. Pinocchio's pupil: using eyetracking and pupil dilation to understand truth telling and deception in sender-receiver games. *Am. Econ. Rev.* 2010, *100*, 984−1007.

35. Berrien F. K., Huntington G. H. An exploratory study of pupillary responses during deception. *J. Exp. Psychol.* 1943, *32*, 443−449.

36. Lubow R. E., Fein O. Pupillary size in response to a visual guilty knowledge test: new technique for the detection of deception. *J. Exp. Psychol. Appl.* 1996, *2*, 164−177.

37. Heilveil I. Deception and pupil size. *J. Clin. Psychol.* 1976, *32*, 675−676.

38. de Gee J. W., Knapen T., Donner T. H. Brein en Cognitie (Psychologie, F.), decision-related pupil dilation reflects upcoming choice and individual bias. *Proc. Natl. Acad. Sci. U. S. A.* 2014, *111*, E618−E625.

39. Grimes M., Valacich J. Mind over mouse: the effect of cognitive load on mouse movement behavior. *Thirty Sixth Int. Conf. Inf. Syst.* 2015, *12*(13).

40. Hibbeln M., Jenkins J. L., Schneider C., Valacich J. S., Weinmann M. Investigating the effect of insurance fraud on mouse usage in human-computer interactions. In *35th International Conference on Information Systems: Building a Better World Through Information Systems, ICIS 2014*; Association for Information Systems, 2014.

41. Dale R., Roche J., Snyder K., McCall R. Exploring action dynamics as an index of paired-associate learning. *PLoS One* 2008, *3*, e1728.

42. Maldonado M., Dunbar E., Chemla E. Mouse tracking as a window into decision making. *Behav. Res. Methods* 2019, *51*, 1085−1101.

43. Monaro M., Gamberini L., Sartori G. The detection of faked identity using unexpected questions and mouse dynamics. *PLoS One* 2017, *12*, e0177851.

44. Chen M. C., Anderson J. R., Sohn M. H. What can a mouse cursor tell us more? Correlation of eye/mouse movements on web browsing. In *CHI'01 Extended Abstracts on Human Factors in Computing Systems*; Association for Computing Machinery: Seattle, Washington, 2001; pp. 281—282.

45. Krishnamurthy G., Majumder N., Poria S., Cambria E. A deep learning approach for multimodal deception detection. *arXiv [cs.CL]* 2018.

46. Abouelenien M., Pérez-Rosas V., Mihalcea R., Burzo M. Deception detection using a multimodal approach. In *Proceedings of the 16th International Conference on Multimodal Interaction*; Association for Computing Machinery: Istanbul, Turkey, 2014; pp. 58—65.

47. MacKenzie I. S. *Human-Computer Interaction: An Empirical Research Perspective*; Morgan Kaufmann Publishers: San Mateo, CA, 2013.

Bionotes

Moritz Maleck
Human-Computer Interaction Group,
University of Bamberg, Bamberg, Germany

Moritz Maleck is a researcher of the Human-Computer Interaction Group at the University of Bamberg. His main research interest is the tracking of user interactions, especially eye and mouse tracking, and its application for special contexts.

Dr. Tom Gross
Human-Computer Interaction Group,
University of Bamberg, Bamberg, Germany
@uni-bamberg.de

Dr. Tom Gross is full professor and founding chair of Human-Computer Interaction at the University of Bamberg, Germany since 1 March 2011. His main research focus is Human-Centred Computing (Human-Computer Interaction HCI, Computer-Supported Cooperative Work, and Ubiquitous Computing). He has been teaching Human-Computer Interaction HCI, Computer-Supported Cooperative Work, and Ubiquitous Computing at various European universities and given keynote talks and tutorials at international conferences. He received his Ph.D. and Habilitation from the University of Linz, Austria.