Detecting the *Doubt Effect* and *Subjective Beliefs* Using Neural Networks and Observers' Pupillary Responses

Xuanying Zhu, Zhenyue Qin, Tom Gedeon(✉), Richard Jones, Md Zakir Hossain, and Sabrina Caldwell

The Australian National University, Canberra, Australia
xuanying.zhu@anu.edu.au, tom@cs.anu.edu.au

Abstract. We investigated the physiological underpinnings to detect the 'doubt effect' – where a presenter's subjective belief in some information has been manipulated. We constructed stimulus videos in which presenters delivered information that in some cases they were led to doubt, but asked to "present anyway". We then showed these stimuli to observers and measured their physiological signals (pupillary responses). Neural networks trained with two statistical features reached a higher accuracy in differentiating the doubt/manipulated-belief compared to the observers' own veracity judgments, which is overall at chance level. We also trained confirmatory neural networks for the predictability of specific stimuli and extracted significant information on those stimulus presenters. We further showed that a semi-unsupervised training regime can use subjective class labels to achieve similar results to using the ground truth labels, opening the door to much wider applicability of these techniques as expensive ground truth labels (provenance) of stimuli data can be replaced by crowd source evaluations (subjective labels). Overall, we showed that neural networks can be used on subjective data, which includes observer perceptions of the doubt felt by the presenters of information. Our ability to detect this *doubt effect* is due to our observers' underlying emotional reactions to what they see, reflected in their physiological signals, and learnt by our neural networks. This kind of technology using physiological signals collected in real time from observers could be used to reflect audience distrust, and perhaps could lead to increased truthfulness in statements presented via the Media.

Keywords: Neural networks · Pupillary responses · Information veracity
Doubt · Trust · Subjective belief · Semi-unsupervised training

1 Introduction

People cooperate, solve problems and enhance social bonds by exchanging information and knowledge. To ensure enduring bonds and to achieve collaborative goals, these communications should be honest and faithful [1] so that people can navigate the information and knowledge with confidence and trust. However, with the proliferation of technologies, information can be easily generated and manipulated. Credibility becomes problematic under the weight of manipulated information such as fake news

L. Cheng et al. (Eds.): ICONIP 2018, LNCS 11304, pp. 610–621, 2018.
https://doi.org/10.1007/978-3-030-04212-7_54

and exaggerated or incorrect advertising, which could potentially cause people being deceived to suffer from grave consequences to their personal lives or administrative decisions. Therefore, from an evolutionary perspective, skills of knowing whom and what to trust is indispensable [2].

Nonetheless, according to Bond and Depaulo [3, 4] when being asked to provide direct veracity judgments of the question such as ‘Is that person lying or telling the truth?’, people are barely better than chance at consciously recognising a dishonesty, with an average accuracy of 54%. This accuracy improves slightly when people’s judgments are assessed indirectly even though they may not be aware that they are being lied to [5]. Additionally, DePaulo et al. [3] and Albrechtsen et al. [6] found that people’s quick and intuitive judgments of dishonesty are more accurate than their slow and deliberative judgments made after conscious reasoning. This is possibly because the unconscious dishonesty detection which results in intuitive judgments may happen at such an early stage that it has not reached consciousness [5]. These findings suggest that people may be able to unconsciously pick up subtle cues from a deception event, advocating the merit of using physiological signals as an unconscious indicator or indirect veracity detector.

Physiological responses maintained by the autonomous nervous system indicate mental state changes such as cognitive load [7] and emotions [8] without conscious awareness [9]. These responses also change with the detection of abnormalities, since unpleasant stimuli induce people’s defensive reactions, marked by anxiety and avoidance which can be assessed with physiological signals [10]. Thus, it seems highly plausible to search for valid physiological indicators of observed dishonesty, which can be considered as an abnormality in observed behaviour as compared to honesty.

Van’t Veer [5] attempted to find a physiological marker of unconscious veracity detection by examining observers’ finger skin temperature, and found that when participants were observing a liar, their finger skin temperature decreased over time. When participants were informed in advance that their goal was to differentiate between liars and truth-tellers, their finger skin temperature declined more when they were watching liars contrasted with watching truth-tellers. Inspired by these findings, in a later study, Van’t Veer [5] explored the use of observers’ pupillary responses in discerning dishonest answers. She demonstrated that dishonesty evoked greater pupillary responses, with pupil size first increasing more and later decreasing more. This result is consistent with [11] in which people were found to have greater eye responses, namely increased fixations and durations of eye gaze, to manipulated areas of images. Another paper has shown that resting heart rate can be correlated with dishonesty detection [12].

Despite physiological correlates of dishonest information being examined previously, we could find no work in which physiological signals from observers were used to identify manipulated subjective belief, in which the presenter is not explicitly intending to deceive. Such *manipulated subjective belief* or *doubt* in information can be considered as a subtle form of deception as the presenters did not mention their doubt, but this is at most a sin of omission not commission.

This paper is organised as follows. Section 2 lists the hypotheses we examine, followed by Sect. 3, which details the experimental design and data collection process. Section 4 describes the feature extraction process and classification models. Results are presented in Sect. 5 and discussed in Sect. 6. Finally, limitations and future work are provided in Sect. 7 before conclusions are drawn in Sect. 8.

2 Hypotheses

Human consciousness is predisposed to trusting the veracity of information, as reported in the literature (see Sect. 1). Thus, when asked to evaluate the veracity of the stimulus content, we expect that participants' judgment will not differentiate between doubting and trusting presentations better than chance. We expect pupillary changes to be significantly different when viewing doubting and trusting conditions, where the reactions to doubting (manipulated subjective belief) are stronger. Thus, a Neural Network classifier trained with pupillary responses should provide a better estimation than human conscious judgments on predicting stimulus veracity.

3 Experimental Design

3.1 Stimuli

Two extracts for popular science books were constructed, phrased and formatted like book publisher advertising materials as Fig. 1 shows. The first was modified from the publisher description of a book "The Salt Fix" written by James DiNicolantonio [13] which describes the benefits of salt in chronic cardiovascular disease. The surname of the author was shortened to Nicol to reduce ethnicity effects. The second extract was summarized and extensively modified from a research paper written by Steven Stanford [14] which presents the feasibility of enzymes in curing diabetes. The author's surname was modified from Stanford to Stafford to avoid the potential celebrity effect of having the same name as a famous university, and a book cover was also constructed. These two materials were chosen because both contents contradict general beliefs that salt is a cause of cardiovascular disease, and that diabetes is not curable. Note that our work is neither a criticism nor an endorsement of their publications.

Four videos were subsequently recorded, each of which consists of an individual presenting one of the above-mentioned book extracts. To construct these videos, one female and one male volunteer from our University were recruited as actors, with Ethics Approval obtained from the Australian National University Human Research Ethics Committee. Neither was a professional actor. After giving their informed consent, they first presented one book extract which they could presume to be true (original subject belief condition). In the brief period before presenting the second extract they were told "Sorry the next one is a bit bogus, please present it anyway". Then, they presented the other extract (manipulated subject belief condition). As we constructed both extracts, this does not reflect our views on the source publications.

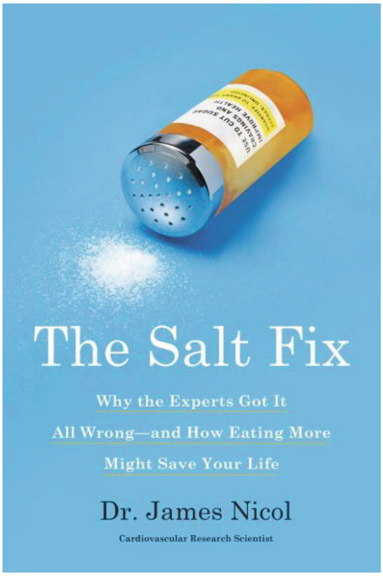
Each extract was read as a belief condition first and a doubt condition second. A camera was placed at 1 m from the actors and filmed them from the chest up. In this way, four videos were recorded, ranging from 27 s to 39 s in length. A laptop was subsequently used to display the videos to the participants during the veracity judgment session.

SALT

For many years we've been told
that salt is just pure bad for you.

Recent research
by leading cardiovascular researcher
Dr James Nicol
upends the low-salt myth,
proving that salt may be a cure
—rather than a cause of,
—our country's chronic disease crises.

Dr. Nicol has shown
how eating the right amount
of this essential mineral
will help you beat sugar cravings,
achieve weight loss,
improve athletic performance,
increase fertility,
and reduce migranes.



DIABETES

For many years we've been told
That diabetes drugs can not cure diabetes.

Recent research
By leading immunology researcher
Dr Steven Stafford
upends the no-cure myth,
proving a cure for diabetes is
—possible, and that diabetes
—is no longer a critical chronic disease.

Dr Stafford has shown
how taking the right enzyme blocker
for an essential amino acid,
will help you control diabetes,
achieve weight loss,
improve athletic performance,
increase fertility,
and reduce eczema.

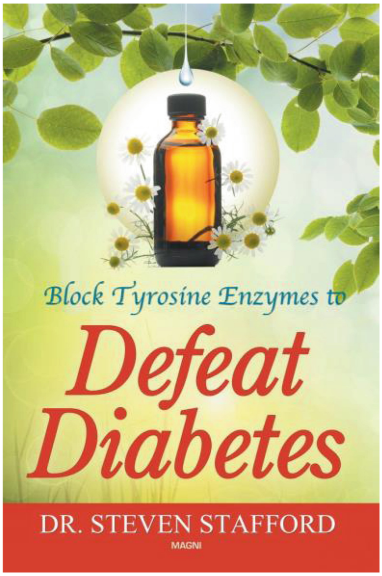


Fig. 1. Two book extracts about salt (above) and diabetes (below)

3.2 Participants

Thirteen students were recruited with Ethics Approval obtained from the Australian National University Human Research Ethics Committee. Five were excluded based on one or more predefined exclusion criteria: for having a history of cardiovascular disease, being acquainted with the stimulus video presenters, and technical failures of the sensors. The final sample consisted of eight participants, four males and four females, from 18 to 24 years in age (Average = 20.6, Standard Deviation = 2.3) with normal or corrected-to-normal vision and hearing.

3.3 Measure and Sensor

Pupillary dilation provides indications of changes in mental states and the strengths of mental activities [15]. Several papers have reported that pupil size constitutes a response to emotionally engaging stimuli in which the pupil is significantly bigger after positively and negatively arousing stimuli than after neutral stimuli [15]. Similar patterns have been found in deception detection where dishonesty evoked greater pupillary responses [5]. In this study, pupillary size was captured using an EyeTribe eye tracker with a sampling rate of 60 Hz [16].

3.4 Procedure

The experiment was conducted with each individual participant in the same quiet experiment room. Participants were forewarned that their goal was to identify the veracity of the presented content. The rationale was: because dishonest information has a higher chance of getting caught when observers are alerted to the possibility of dishonesty in advance [5], then this may also apply to manipulated (doubting) beliefs. They provided written informed consent, and filled in a questionnaire to collect demographic and health characteristics that may affect pupillary responses. Eye gaze calibration for the tracker was performed. Participants watched the two videos one by one and were asked to provide responses to a question of “To what extent do you believe in the content?” on a five-item Likert scale (from 1 = strongly no, to 5 = strongly yes) with a confidence level (from 0% to 100%). The videos were presented in an order balanced way to avoid effects of presentation ordering.

4 Methodology

After data collection, eight manipulated and eight unmanipulated subjective belief observations were obtained. In the rest of this paper we will mostly use *doubting* and *trusting* to succinctly express “manipulated subjective belief” and “unmanipulated subjective belief”, respectively. Two matching observations were excluded from further analysis due to intermittent sensor data loss. This results in a total of twelve complete responses, including their conscious veracity judgments and physiological sensors recordings.

4.1 Feature Extraction

The raw pupillary data was processed to obtain features. The pupillary responses were calculated from each observer's pupil size acquired in viewing each video. For each second, the EyeTribe returned 60 frames of raw pupillary data and therefore, for example, a 27-second video would result in 1,620 frames of data recorded. The pupillary data were extracted from the beginning of the video playing to the end of the video. The extracted pupil size data were subsequently normalised across all videos viewed by each observer to reduce the effect of individual bias due to the naturally varying pupil sizes among participants, since the significant signal is the variation in the pupil size for an individual, not the magnitude of the pupil size.

Linear interpolation was applied to missing pupil size data caused by occasional eye blinks. This procedure was employed on the pupil data of left and right eyes separately. The interpolated pupil size values were averaged to obtain a single pupillary response signal for each participant, and their minimum and mean of processed pupillary data during each video watching session were extracted as features.

4.2 Model Description

An ensemble of five artificial neural networks were trained on observers' pupillary responses to predict presenters' subjective beliefs individually and jointly as well as observers' conscious veracity judgments. All neural networks had a sigmoid hidden layer of size 100 and an output layer with two output neurons. They were trained with the Adam optimizer [17] without any adaption using backpropagation with the Cross-Entropy loss function.

The most common method for cross-validation is k-fold which randomly partitions data into k equal sized subsamples and uses one subsample as validation data, and the remaining k-1 as training data. However, for human data, a continuous segment of physiological data with more than one data point can reflect a human's responses to a stimulus. Training a classification model on random splits of data is not adequate, unless all data from one human is guaranteed to be within either the training set or the testing set for each run. This method of leaving all data for one human out is called leave-one-participant-out, which was used in this study. Pupillary data from one observer was used as the testing set, and those from the remaining participants formed the training set, and repeated for all, averaging to calculate the final result reported.

5 Results

5.1 Veracity Judgments

The average accuracy of observers' conscious veracity judgments (see Table 1) was 50% with standard deviation of 0.43, which is at the prima-facie chance level of 50%, because there were two options for the observers over balanced numbers of video stimuli. The average accuracy of consciously identifying doubting presentations was 40%, lower than that of identifying trusting presentations which was at 60%. This could imply that people are better at identifying trusted information compared with doubted content. The accuracies of predicting veracity of subjective beliefs vary greatly between presenters, at the rate of 80% and 20% respectively, which may indicate that the veracity of behaviours from one presenter was easier to identify compared with those from the other. The 80% represents well over chance recognition of doubting for one presenter, while the 20% represents well *below* chance for the other, indicating our observers were consistently wrong on that presenter.

Table 1. Accuracies of doubting and trusting prediction from observers' verbal responses

	By manipulation			By presenters	
	Doubt condition	Trust condition	Total	Presenter 1	Presenter 2
Verbal responses	40%	60%	50%	80%	20%

5.2 Pupillary Responses

Figure 2 displays observers' pupil diameter when they were watching doubting and trusting presentations. Visually, regardless of the veracity of the content, there is an overall wave-like changing pattern that may reflect pupillary responses for a common process while viewing a video stimulus.

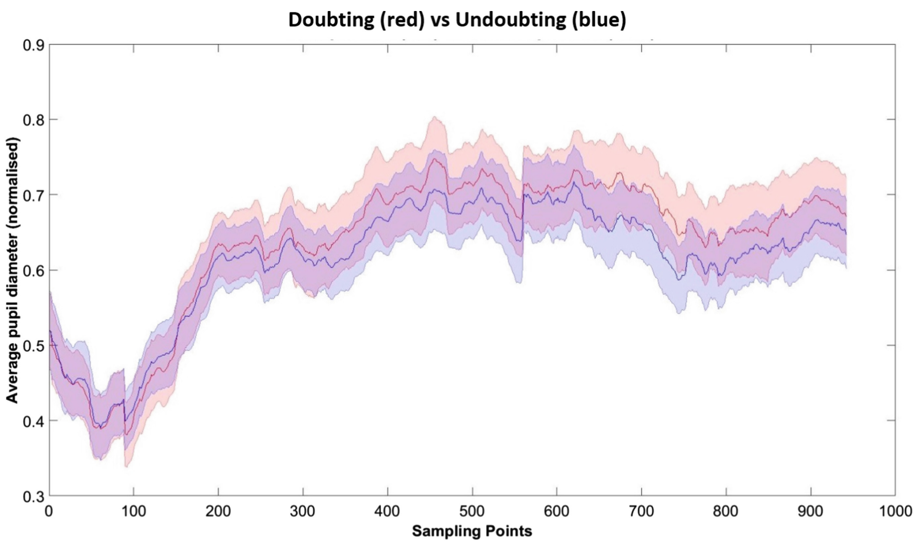


Fig. 2. Average pupillary response to doubting and trusting presentations over the time

The responses to doubting versus trusting presentations started with similar patterns and then diverged at around frame 90. Such divergence continued until at around frame 190 when the pupil diameter changed to opposite directions for doubting versus trusting presentations. A third divergence happened at around frame 560 but the directions for doubting versus trusting stimuli viewing remained the same. The first two diverging frames may indicate the beginning of human unconscious discrimination of stimuli content.

By observation, over time, when viewing a stimulus with doubted beliefs, observers' pupil diameters had greater changes, first decreasing more and then increasing more. To test whether the pattern of pupillary diameters was statistically different when observing a doubted belief compared to a trusting belief, a two-sample Kolmogorov-Smirnov test was performed. The result showed that the pupil dilated differently with a statistical significance ($p < 0.001$) for doubting stimuli observations compared with trusting stimuli observations. This provides a clear indication that the pupillary response can be of use for computational classification.

5.3 Relationship Between Pupillary Responses and Veracity Judgments

To assess whether observers' pupillary responses are predictive of their veracity judgments, the first neural network in the ensemble was trained with observers' pupillary features to predict their veracity judgments for all observers across all videos. The average accuracy was 75.8%, indicating that people's pupillary responses are positively correlated with their conscious veracity judgments.

5.4 Relationship Between Pupillary Size, Doubt Effect and Subjective Beliefs

As described above, doubting presentations resulted in greater observers' pupil arousal responses with a high statistical significance. To explore the relationship between observers' pupillary responses and presenters' doubt effects, the second neural network was trained with features extracted from observers' pupillary size to predict whether presenters doubted the presented content. From Table 2, the final accuracy for all observers across all videos was 58.3%, which is higher than the chance level of 50%. This accuracy is also higher than that of observers' conscious veracity judgments, implying that in general, neural network classifiers trained using pupillary responses can better identify doubting presentations than human conscious judgments.

Table 2. Accuracies of doubting and trusting prediction from observers' pupillary responses

	By manipulation			By presenters	
	Doubt condition (Specificity)	Trust condition (sensitivity)	Total (accuracy)	Presenter 1	Presenter 2
Pupillary responses	50%	66.7%	58.3%	100%	33%

The average accuracy of identifying doubting presentations with pupillary responses was 50%, lower than that of identifying trusting presentations which was 66.7% (see Table 2). This shows neural networks trained on human pupillary responses are also better at identifying trusting than doubting content. Also, these classifiers are more accurate, as their accuracies are *both* higher than those of observer's conscious veracity judgments. Besides, despite the improvement of accuracies, the prediction of pupillary responses is still not completely accurate. While this could be due to the limited features extracted from pupillary responses, this could also imply that at the physiological level, observers could still be fooled by presenters' behaviours.

To further explore whether different presenters' behaviours can be discerned from observers' physiological data, a neural network was trained for *each* presenter with pupillary responses from observers when they viewed that presenter. As shown in Table 2, the accuracies of predicting veracity of subjective beliefs with pupillary responses vary greatly between presenters, from 100% to 33%. These accuracies are both higher than those from observers' verbal responses which were 80% to 20%. While this shows the superior prediction of pupillary responses over verbal responses, it also reveals the possible stimuli bias introduced by different presenters. The classifier for Presenter 1 improves the prediction to perfect, however, while the accuracy of that for Presenter 2 is improved, but is still below chance level. This could be owing to the different behaviours by the two presenters, with one being an obvious disbeliever and another being either a good dissembler, or who was better at following the instructions literally, and tried sufficiently hard to "present it anyway" to be more plausible on the doubting condition than on the trusting condition.

5.5 Relationship Between Pupillary Size and Veracity by Voting

The last neural network was trained on observers' pupillary responses with labels derived from their verbal responses, with voting between subjects. That is, the majority classification is used. Leave-one-stimulus-out cross validation was used, meaning that responses to all but one videos were used as training data, and those to one video used as testing data. The accuracy was 75% on average. This is a "semi-unsupervised" learning as the real labels have not been used. Instead, the labels are "crowdsourced" from the experiment participants' verbal responses.

6 Discussion

With this preliminary study, we explored people's automatic and non-conscious pupillary responses to observing a doubting presentation, a subtle form of dishonest information, as well as their more conscious verbal judgments of the veracity in a presenter's video. For people's ability to accurately indicate whether the provided information is valid, it was predicted that conscious veracity judgment would not detect doubting much better than chance. Consistent with (but slightly worse than) earlier findings about the accuracy of people's conscious judgments on the veracity of smiles [18] and anger [19], observers in this study were found to be accurate about 50% of the time, a performance that was not different from chance. Our analysis showed that the two presenters in our study consist of an obvious doubter and a good dissembler; this could still imply that human conscious judgments on veracity may not be better than chance in general in evaluating direct deception and doubting. Future research could explore conscious veracity judgments on more complex emotions or even more subtle deception activities.

Observers' pupillary size was measured during the viewing of doubting stimuli. Different from the previous findings where pupillary size was discovered to first increase and then decrease while watching disbelief [5], the pupillary changes in this study started with an initial short decrease followed by a longer increase and ended with a minor decrease. Consistent with previous findings [20, 21], in this study the initial decrease in pupil dilation lasted slightly longer than 1 s, possibly due to the light reflex related to the changes in environmental illumination caused by the playing of stimulus. The subsequent increase may reflect the preparation for the human body to react appropriately to external stimuli. Since pupil dilation occurs with temporal changes of visual attention and is found to reflect personal preferences [22, 23], this reaction thus addresses the importance and preference for what is being observed, and therefore may assist in estimating the trustworthiness of others.

As for the pupillary responses to doubting and trusting content, it was found that observers' pupil diameters had greater changes in the doubting condition, consistent with earlier findings in the literature indicating that dishonesty evokes greater pupillary responses (on more obvious deception behaviours) [5]. Our result further suggests that dishonesty may be unconsciously perceived, upholding the evolutionary view that accurate detection of obvious deception and subtle doubt is adaptive and should be advantaged by natural selection [1].

With regard to the comparison between conscious judgments and a computational classifier trained on unconscious pupillary responses for identifying the veracity of presented information, better estimation was achieved by the classifier. We also found that people's pupillary size was positively correlated with their conscious judgments. Taken together, the results provide evidence that although humans cannot consciously discriminate doubted information from trusted information correctly, better ability of detecting doubted information may be present at an unconscious level and this ability can be accessed by computational classifiers. The superior veracity detecting ability of human unconscious pupillary responses over conscious judgments also occur in other areas such as estimating realness of two basic emotions [18, 19]. This suggests that while unconscious responses from human instinctive ability, which has been adaptively evolved by natural selection, can make efficient and effective use of deception cues, the resulting conscious judgments are made inaccurate under the influence of conscious biases and social rules [1]. Further exploration can be performed to examine the applicability of human unconscious responses to the veracity of information or emotions with subtler but perhaps socially important differences.

7 Limitations and Future Work

The stimuli videos used in this study could be increased in number, though the results were highly statistically significant, at the $p < 0.001$ level. Also, we assumed that presenters made the same level of effort in hiding their subjective beliefs, and that the beliefs were synchronised with the experimenter instructions. That is, when they were presented the content of a book extract, they would believe in it unless doubt was suggested. In a future study, more actors should be recruited, and their acting effort as well as their own beliefs should be collected by a structured interview after their video performance is recorded. Stronger conclusions may also be able to be drawn in subsequent studies with more observers. More statistical features from pupil responses could be investigated to see if this leads to more accurate computational classifiers. Finally, observers' other physiological signals such as galvanic skin response, skin temperature and blood volume pulse could also be investigated. With increasing amounts of data collected from wider groups of participants, recent deep learning models, such as Recurrent Neural Networks or Long Short-Term Memory models, can potentially be trained to achieve more accurate results.

8 Conclusion

Our work explored physiological signals to detect the 'doubt effect' where a presenter's subjective belief in some information was manipulated. When doubted information was observed, discernible physiological indicators are present: following an initial light reflex, pupillary changes fluctuated more with an increase and then a decrease. Neural network ensembles trained with 2 pupillary features can reach a higher accuracy in differentiating doubting and trusting information compared with the same observers' conscious veracity judgments. We demonstrated that neural networks trained within-stimulus can provide insights into the dataset. Human data is complex, and this approach may be useful more generally. We also introduced a form of semi-unsupervised training, using crowd-sourced labels rather than the ground-truth.

References

1. ten Brinke, L., Vohs, K.D., Carney, D.R.: Can ordinary people detect deception after all? Trends Cogn. Sci. **20**, 579–588 (2016)
2. Von Hippel, W., Trivers, R.: The evolution and psychology of self-deception. Behav. Brain Sci. **34**, 1–16 (2011)
3. Bond Jr., C.F., DePaulo, B.M.: Accuracy of deception judgments. Pers. Soc. Psychol. Rev. **10**, 214–234 (2006)
4. DePaulo, B.M., Bond Jr., C.F.: Beyond accuracy: bigger, broader ways to think about deceit. J. Appl. Res. Mem. Cogn. **1**, 120–121 (2012)
5. van't Veer, A.: Effortless morality: cognitive and affective processes in deception and its detection. Dissertation, Tilburg. University (2016)
6. Albrechtsen, J.S., Meissner, C.A., Susa, K.J.: Can intuition improve deception detection performance? J. Exp. Soc. Psychol. **45**, 1052–1055 (2009)
7. Chow, C., Gedeon, T.: Classifying document categories based on physiological measures of analyst responses. In: 2015 6th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), pp. 421–425 (2015)
8. Kreibig, S.D.: Autonomic nervous system activity in emotion: a review. Biol. Psychol. **84**, 14–41 (2010)

9. Tomaka, J., Blascovich, J., Kelsey, R.M., Leitten, C.L.: Subjective, physiological, and behavioral effects of threat and challenge appraisal. J. Pers. Soc. Psychol. **65**, 248 (1993)
10. Sleegers, W., Proulx, T.: The comfort of approach: self-soothing effects of behavioral approach in response to meaning violations. Front. Psychol. **5**, 1568 (2015)
11. Caldwell, S., Gedeon, T., Jones, R., Copeland, L.: Imperfect understandings: a grounded theory and eye gaze investigation of human perceptions of manipulated and unmanipulated digital images. In: Proceedings of the World Congress on Electrical Engineering and Computer Systems and Science (2015)
12. Duran, G., Tapiero, I., Michael, G.A.: Resting heart rate: a physiological predicator of lie detection ability. Physiol. Behav. **186**, 10–15 (2018)
13. DiNicolantonio, J.: The Salt Fix. Harmony (2017)
14. Stanford, S.M., et al.: Diabetes reversal by inhibition of the low-molecular-weight tyrosine phosphatase. Nat. Chem. Biol. **13**, 624 (2017)
15. Laeng, B., Sirois, S., Gredebäck, G.: Pupillometry: a window to the preconscious? Perspect. Psychol. Sci. **7**, 18–27 (2012)
16. TheEyeTribe: The EyeTribe. http://theeyetribe.com/theeyetribe.com/about/index.html
17. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv Preprint arXiv1412.6980 (2014)
18. Hossain, M.Z., Gedeon, T.: Classifying posed and real smiles from observers' peripheral physiology. In: 11th International Conference on Pervasive Computing Technologies for Healthcare (2017)
19. Chen, L., Gedeon, T., Hossain, M.Z., Caldwell, S.: Are you really angry?: detecting emotion veracity as a proposed tool for interaction. In: Proceedings of the 29th Australian Conference on Computer-Human Interaction, pp. 412–416 (2017)
20. Bradley, M.M., Miccoli, L., Escrig, M.A., Lang, P.J.: The pupil as a measure of emotional arousal and autonomic activation. Psychophysiology **45**, 602–607 (2008)
21. Beatty, J., Lucero-Wagoner, B., et al.: The pupillary system. In: Handbook Psychophysiology (2rd edn.) (2000)
22. Wiseman, R., Watt, C.: Judging a book by its cover: the unconscious influence of pupil size on consumer choice. Perception **39**, 1417–1419 (2010)
23. Gründl, M., Knoll, S., Eisenmann-Klein, M., Prantl, L.: The blue-eyes stereotype: do eye color, pupil diameter, and scleral color affect attractiveness? Aesthetic Plast. Surg. **36**, 234–240 (2012)