

Inferring Others' Hidden Thoughts: Smart Guesses in a Low Diagnostic World

CHRIS N. H. STREET^{1*}, WALTER F. BISCHOF^{2,3}, MIGUEL A. VADILLO⁴ and ALAN KINGSTONE³

¹University of Huddersfield, Huddersfield, UK

²University of Alberta, Edmonton, Canada

³University of British Columbia, Vancouver, Canada

⁴King's College London, London, UK

ABSTRACT

People are biased toward believing that what others say is what they truly think. This effect, known as the truth bias, has often been characterized as a judgmental error that impedes accuracy. We consider an alternative view: that it reflects the use of contextual information to make the best guess when the currently available information has low diagnosticity. Participants learnt the diagnostic value of four cues, which were present during truthful statements between 20% and 80% of the time. Afterwards, participants were given contextual information: either that most people would lie or that most would tell the truth. We found that people were biased in the direction of the context information when the individuating behavioral cues were nondiagnostic. As the individuating cues became more diagnostic, context had less to no effect. We conclude that more general context information is used to make an informed judgment when other individuating cues are absent. That is, the truth bias reflects a smart guess in a low diagnostic world. Copyright © 2015 John Wiley & Sons, Ltd.

KEY WORDS adaptive decision making; adaptive lie detector; context; lie detection; truth bias; truth-default theory

People make for poor lie detectors. They hit an accuracy rate comparable with a coin toss, only marginally above chance (Bond & DePaulo, 2006). That might in part be because they come with a set of systematic biases that can reduce accuracy (Burgoon & Buller, 1994; Gilbert, 1991; O'Sullivan, 2003; Vrij, 2008), such as being biased toward believing that others are telling the truth more often than they actually are (Bond & DePaulo, 2006; McCornack & Parks, 1986). This article argues that the pessimistic view of lie detection is outdated and misguided: Instead, we argue that people make smart judgments from the unreliable information available to them. Specifically, it is argued that the systematic biases are not sources of error but are actually markers of a smart system making informed judgments.

Although people are thought to be poor lie detectors, this stands in stark contrast to research showing that people are skilled at understanding what others are thinking. They have successful strategies to understand internal thoughts, drawing from even the subtlest of clues (e.g., from allusions in speech to eye direction: Clark, 1996; Clark, Schreuder, & Buttrick, 1983; Tomasello, 1995; but see Heyes, 2014, for strategies that give only the illusion of representing others' minds). These work only because the two parties choose to communicate with each other, producing behaviors that indicate their true thoughts. When one party wants to conceal what he or she truly thinks, these cues are not produced or are at least well-hidden. Accuracy is capped by the fact that speakers have good control over their behavior and do not give clear signs to their concealments and deceptions (DePaulo et al., 2003; Sporer & Schwandt, 2006, 2007). So we must ask how it is that people make any sort of social

judgment when the immediately available cues have low diagnostic value. Put another way, how do raters deal with their uncertainty in order to reach a judgment?

The adaptive lie detector

Street (2015) and Street and Richardson (2015) proposed that raters attempt to deal with this type of uncertainty by relying on general context-relevant knowledge to make the most informed guess. Because context information tends to suggest that people are more likely to tell the truth than lie (e.g., DePaulo, Kashy, Kirkendol, Wyer, & Epstein, 1996; Fiedler, Armbruster, Nickel, Walther, & Asbeck, 1996; Grice, 1975; Sperber et al., 2010), raters make an informed guess from this context information and judge others as telling the truth (Street & Richardson, 2014, 2015). Lay people are typically truth biased (Bond & DePaulo, 2006), but this is functional because most people tell the truth (Levine, 2014; Street & Richardson, 2014). Communication only functions effectively if people are assumed to be telling the truth (Grice, 1975), at least most of the time (Sperber et al., 2010; Sperber, 2013). In fact, people do usually tell the truth (DePaulo et al., 1996), with approximately half of all lies told by a very small minority of people (Halevy et al., 2013). When unsure, it is smart to guess that people are telling the truth because most of the statements we hear are truths. In that sense, the truth bias (Bond & DePaulo, 2006; McCornack & Parks, 1986) might reflect an active and functional strategy to make a lie–truth judgment that can *improve* accuracy, rather than being an erroneous bias that *impedes* accuracy. We put this account to the test here.

A bias that results from an active strategy to deal with low diagnosticity in the environment must be flexible and adaptive

*Correspondence to: Chris N. H. Street, International Research Centre for Investigative Psychology, Ramsden Building, University of Huddersfield, Queensgate, Huddersfield, HD1 3DH, UK. E-mail: cstree1986@gmail.com

to the current situation. That is, people should not default to a truth-biased position (Gilbert, 1991; Levine, 2014; Mandelbaum, 2014). Rather, they should be biased in the direction of both their prior experience with and their current understanding of the situation they find themselves in (Blair, 2006; Brunswik, 1952; Masip, Alonso, Garrido, & Herrero, 2009; Street & Richardson, 2014). This should apply particularly when the available cues in the immediate environment are weak and lacking in diagnosticity (Garcia-Retamero & Rieskamp, 2008, 2009; Jekel, Glöckner, Bröder, & Maydych, 2014).

Consistent with the hypothesis that people use flexible and adaptive strategies to detect lies, there is evidence that the bias is not steadfastly skewed toward believing (cf. Gilbert, 1991). A bias toward disbelieving is observed when the context suggests that most people will lie (Street & Richardson, 2015; DePaulo & DePaulo, 1989; Levine, Park, & McCornack, 1999; Millar & Millar, 1997). Police officers typically demonstrate a lie bias (Meissner & Kassin, 2002), and they report expecting most people to lie to them (Moston, Stephenson, & Williamson, 1992). Training can also cause a shift from a truth-biased to a lie-biased position (Blair, 2006; Masip et al., 2009). The direction of the bias seems to adapt to the raters' understanding of the current context.

This adaptive lie detector position (Street, 2015; Street & Richardson, 2015) claims that context is relied upon more heavily when the more individuating cues in the immediate environment (behavioral cues) have low diagnosticity. When more individuating cues are diagnostic, people rely more on those cues (Brunswik, 1952; cf. Bond, Howard, Hutchison, & Masip, 2013). Brunswik (1952) introduced the term "vicarious functioning" to note that behavior is purposive and goal directed. To meet that goal, people make more or less use of different pieces of information depending on what is currently available in the environment and on what is more useful to make a judgment. That is, cognition *adapts* to the demands of the environment and is *adaptive* or functional in a given situation (Simon, 1990). Might lie-truth judgments show just such adaptivity? And is the truth bias, often characterized as an error in the process (Burgooon & Buller, 1994; Gilbert, 1991; O'Sullivan, 2003; Vrij, 2008), really a marker of a smart, context-informed "guess"? If the bias is an adaptive strategy to deal with uncertainty, it should be observed only in the absence of more useful individuating information—whether that is because the information is lacking or because the present information has low diagnosticity (Gigerenzer, Hertwig, & Pachur, 2011). By making use of more generalized (and necessarily simplified) context information, people can reach satisfactory decisions from the limited information provided to them (Brunswik, 1952; Simon, 1990; see also Wolf, 1999, cited by Degani, Shafto, & Kirlik, 2006).

For instance, when negative information is more likely to be missing than positive information, it is reasonable to guess that the missing information is more likely to be negative (Garcia-Retamero & Rieskamp, 2008, 2009). Indeed, individuals make adaptive inferences about what missing information entails, given the current context, and use that to

inform their decision making (Jekel et al., 2014; Meiser, Sattler, & von Hecker, 2007; Nadarevic & Erdfelder, 2013). Similarly, when people have very little information about someone, they "fill in" their understanding of the person with more generic stereotypic knowledge (Kunda, Davies, Adams, & Spencer, 2002). But as more individuating information becomes available, the stereotypic information is replaced with the more diagnostic, individuating information about the person (Kunda et al., 2002; Woolley & Ghossainy, 2013; Yzerbyt, Schadron, Leyens, & Rocher, 1994). People may not trust "the politician," but when they learn more about Julie who lives on the next street, enjoys sports, and volunteers on the weekend, they will displace the more generic stereotyped view with an informed view. Also, people rely less on base rate information, which gives overarching information about the context, when the more immediately available individuating cues have greater diagnosticity (Ginossar & Trope, 1980; Koehler, 1996). That is, base rates are used only when they can contribute something informative to the judgment (Bar-Hillel, 1990).

In the current study, we explore whether a similar process takes place in the case of lie detection. Specifically, we ask whether raters make informed judgments by using context when the immediately available information has low diagnosticity. We predict that when the available information is diagnostic, raters make use of it to inform their judgment. But when that information reduces in diagnosticity, people will switch to using more simplified rules that are contingent on the current context.

METHODS

Participants

Eighty participants took part in return for \$5. The computer crashed mid-experiment while running one participant and another participant needed to leave before completing the experiment, and so these two participants were excluded from analysis. In addition, the age of one participant and both the age and sex of another participant were not recorded, but these participants were not removed from analysis. For the participants where age and/or sex were recorded, the mean age was 23.00 years ($SD=5.03$, range: 17–44), 57 of whom were female.

Procedure and materials

There were three parts to this study: (i) a training phase that taught participants how diagnostic four behavioral cues were in detecting liars in a trivia game; (ii) a context manipulation that informed participants the majority of trivia game players lied or told the truth; and (iii) a test phase that assessed whether context information is weighted more heavily in the judgment as the behavioral cues become less diagnostic.

The training task was modeled on that of Bröder (2000). In this task, participants were told that a large-scale study had been previously conducted in which over 450 people

took part. In reality, there was no such earlier study. We told participants that people in that study took part in a trivia game and that there was an opportunity to cheat while the experimenter was out of the room. Each of the trivia game players was supposedly interviewed at the end of the game, with one question that specifically asked whether they had cheated. Although all of the trivia game players supposedly denied cheating, some denied it by telling the truth (i.e., they actually did not cheat) while others denied it by lying (i.e., they cheated and later denied cheating). That is, if the trivia game players cheated, they must have lied later.

The interviews with the fake trivia game players were supposedly coded for four behaviors: (i) pitch of voice; (ii) degree of facial emotional expressivity; (iii) the number of silent periods in the middle of sentences; and (iv) the number of self references such as “I” and “me.” Rater participants in the current study were told that they were to learn how the presence of each of these behaviors related to whether the person was lying or telling the truth. That is, they were to be trained on the diagnosticity of each behavior. An example of a training trial is shown in Figure 1. The cue shown at the top of the screen was supposedly expressed by the player in the trivia game, and the rater had to judge whether the trivia game player was lying or telling the truth.

Because the trivia game did not really take place, the diagnosticity of the four cues could be controlled and manipulated. For each participant, we randomly selected without replacement four different diagnosticity values, one for each cue, on condition that this would result in an equal presentation of lies and truth. These values were an indication of how often people told the truth when that cue was *present*—participants never learnt the diagnosticity of the absence of a cue. The diagnosticity values were 20%, 30%, 40%, 50%, 60%, 70%, and 80%. For example, for one participant, these values could be 20, 50, 60, and 70; for another participant, they could be 30, 40, 50, and 80.

Each cue was presented for a block of 40 trials in the training phase, with the cue order counterbalanced across participants. Participants judged whether a presented behavior indicated that

the speaker lied or told the truth. They were then told if they guessed right or wrong. A bar chart kept track of how many trivia players had supposedly lied when they had shown this cue and how many had told the truth. After each block, there was a 30-second break. To ensure that the base rate of honesty was equal to the base rate of deception, participants judged an equal number of truthful and deceptive trivia game players during training.

Each trivia game player was assigned a participant number, as can be seen in Figure 1. A participant number was never seen more than once throughout the experiment (i.e., across the training and test phases).

Before the start of the test phase, participants were shown one of two different pieces of information, which defined the context manipulation. Critically, the context information was not given during the training phase so that the learning of the cue diagnosticities was unaffected by the context manipulation. In one condition, participants were told that the trivia game was difficult, and if the trivia game players did as well as they had claimed (and so achieved a cash prize), most people had to cheat and therefore were lying when they denied cheating. In the other condition, participants were told that the trivia game was easy. Most trivia game players could achieve the accuracy level they claimed without having to cheat, and so most people told the truth when they denied cheating. It is important to remember that all trivia game players in the stimulus set are described as having denied cheating. Additionally, participants were told that they would only see a small sample of all the 450 participants.

After the context manipulation, the test phase began. This was identical to the training phase with three important changes. First, there was no feedback after making a judgment to minimize additional learning. Second, for the same reason, the bar graphs were no longer presented. Third, the cues were not blocked: Each trial was randomized and could portray any one of the four cues. There were two blocks of 40 trials, with a 30-second break between the blocks. Participants were fully debriefed. Our institutional review board approved the study.

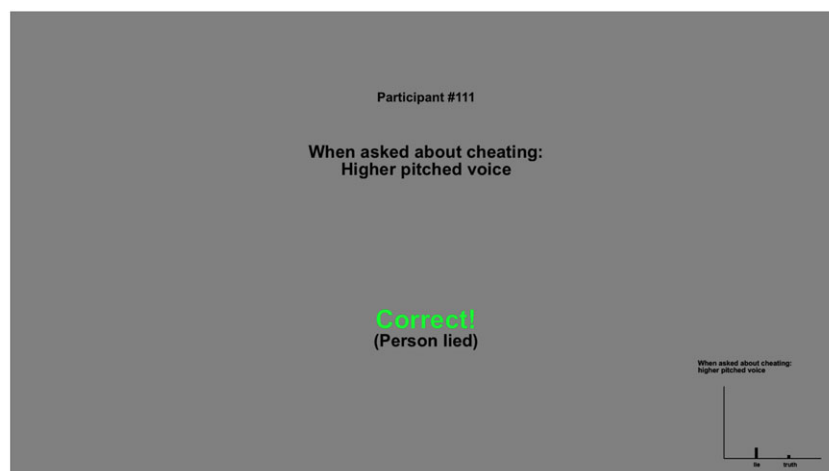


Figure 1. A screenshot taken from the training phase. The cue, “higher pitched voice,” was present on this trial, and participants had to indicate whether they believed that the speaker lied or told the truth. Feedback was given (“correct,” in this instance), and a bar chart tracked the proportion of trivia game players so far who had lied or told the truth when displaying this same cue. Training on each of the four cues was blocked

RESULTS

The context (the trivia game was either easy or difficult, leading to little or mostly lying, respectively) was manipulated between subjects. How diagnostic the cue was of honesty was a second independent variable, ranging between 20% and 80% diagnostic of honesty in steps of 10%. Because there were seven levels of diagnosticity (20–80%) and only four cues, a generalized linear mixed-effects model (GLMEM) was fitted to the data with random effects for participant number. The dependent variable was the proportion of truth judgments (PTJ), calculated as the number of truth judgments divided by the total number of judgments made.

To check that participants understood our context manipulation, we asked them to report what percentage of trivia game players in the test phase had cheated. Four participants did not answer this question. The percentage of cheaters was judged to be lower by those participants in the easy context condition ($M=44.5\%$, $SD=20.69$), suggestive of honest behavior, than in the difficult context condition ($M=58.5\%$, $SD=17.45$) with a medium to large effect size, $t(74)=3.19$, $p=.002$, $d=.73$.

The main results of the experiment are presented in Figure 2. A 2 (context: easy or difficult trivia game, between subjects) \times 7 (cue diagnosticity: 20–80% diagnostic of honesty, within subjects) GLMEM with random effects fitted for participants found a main effect of cue diagnosticity, $F(6, 277.81)=41.03$, $p<.001$,¹ showing an increasing PTJ as the cue became more diagnostic of honesty; see Figure 2. In short, these data indicate that the training phase had a coherent and significant effect on performance. There was also a main effect of context, $F(1, 71.85)=24.29$, $p<.001$: More truth judgments were made when the trivia game was thought to be easy (and so trivia game players truthfully denied cheating: $M=.57$, $SD=.21$, 95% confidence intervals [.52, .62]) than when they thought that the game was difficult ($M=.41$, $SD=.21$, [.36, .46]).

Critically, there was a cue diagnosticity \times context interaction, $F(6, 277.81)=3.28$, $p=.004$. This interaction is illustrated in Figure 2. Figure 2a shows that as the cue diagnosticity increases, that is, as the cue value moves away from 50% in either direction, the effect of context becomes smaller. The difference between contexts is plotted in Figure 2b by subtracting the difficult from the easy context. It is possible that there is greater variance when the cue is nondiagnostic (50%) than when it is highly diagnostic (20% or 80%). Therefore, to control for within-subject variance, the difference values were converted to Cohen's d values to standardize them by the pooled standard deviation. However, the same results are found when using the unstandardized difference values, available on request. A clear quadratic effect can be observed, with an r^2 of .826 (adjusted $r^2=.739$). As the cue diagnosticities become weaker, that is, as we approach 50%, the use of context has a greater effect.

¹Note that the effect sizes for GLMEM analyses are not reported because there is no agreement on how best to decompose the variance from random effects (e.g., Nakagawa & Schielzeth, 2010).

Additionally, we asked whether participants were biased in the direction of the context information when the available cues were nondiagnostic (i.e., 50% diagnostic). Two one-sample t -tests found that raters in the difficult context condition were biased toward making more lie than truth judgments ($M=.41$, $SD=.25$, 95% confidence intervals [.33, .48]), $t(39)=-2.38$, $p=.022$, $d=-.36$, while those in the easy condition were truth biased ($M=.67$, $SD=.25$, [.59, .75]), $t(39)=4.36$, $p<.001$, $d=.68$. In short, in the absence of diagnostic information, people rely on context information, and this can cause a lie or truth bias, depending on the context.

Computational model

The behavioral data give the impression that people use both context and individuating cues in their judgments but to different extents so that one piece of information compensates for the lack of diagnosticity of the other (known as a compensatory strategy). But there is a viable alternative explanation: It may be that people use either the context or the individuating cue, depending on which is the more informative (known as a noncompensatory strategy). This latter strategy is taken up by the take-the-best account (Gigerenzer & Goldstein, 1996), which proposes that people use only the most satisficing or “good enough” information that distinguishes two options. To fit our data, such a model would claim that people use only individuating cue diagnosticity information, ignoring context, until it becomes sufficiently lacking in diagnosticity. At that point, people switch to using context information to make their judgment, ignoring individuating cues. We contrasted these two strategies using computational modeling.

Aggregate fit to the data

The parameters of our model were optimized based on aggregated data for each cell of the 2 (context) \times 7 (truth diagnosticity) design. Nelder–Mead simplex optimization was used to estimate the free parameter of each model. D'Errico's (2012) `fminsearchbnd` for MATLAB was used to constrain the bounds of the parameter estimates between .50 and 1.00. We used 10 000 starting parameters randomly selected from within these bounds.

The models assumed that participants learnt the truth-indicative cue values accurately during training. These cue values, denoted V_{CUE} , ranged between 20% and 80% indicative of honesty. We did not give participants a quantitative value for how informative or “diagnostic” the context information would be. Rather, we gave them a qualitative description—“most people cheated.” To quantitatively capture how informative the context information was, this was fitted as the free parameter I_{CTX} in both models.

A naïve Bayes compensatory model integrated the context-general information with the individuating cue values. Context information, or the probability that people

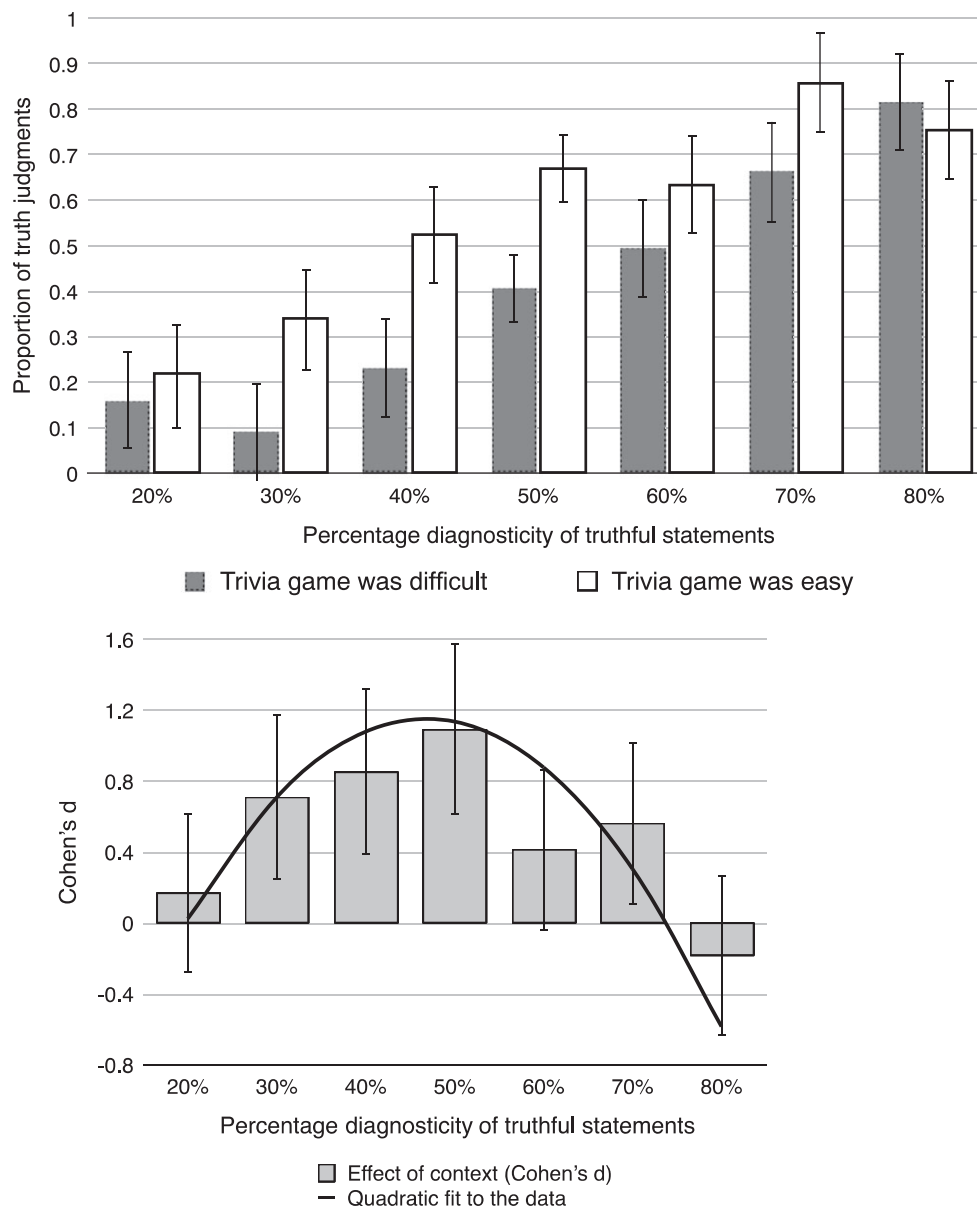


Figure 2. (a, top) The proportion of truth judgments as a function of cue diagnosticity and context. (b, bottom) The difference in proportion of truth judgments for the hard versus easy game context, divided by the pooled standard deviation (i.e., Cohen's d) for each cue. As can be seen, context has greater effect toward 50%. Error bars denote 95% confidence intervals

are telling the truth (CTX), is taken as the likelihood. Individuating cue values in the current context, or $p(V_{CUE}|CTX)$, are the evidence.

$$PTJ = \frac{p(V_{CUE}|CTX) \times CTX}{(p(V_{CUE}|CTX) \times CTX) + (p(V_{CUE}|1 - CTX) \times (1 - CTX))} \quad (1)$$

Or in other words, the PTJ is the result of multiplying the cue value (in this context) with the informativeness of context, which is then standardized by the overall probability of the cue. But the probability that people are telling the truth (i.e., CTX , the context information) depends on whether participants were told that the trivia game was easy or difficult.

To account for this dependence of the value of CTX , the following rule was applied:

$$CTX = \begin{cases} I_{CTX} & \text{trivia game is easy} \\ 1 - I_{CTX} & \text{trivia game is difficult} \end{cases} \quad (2)$$

This allowed us to fit only a single free parameter called the informativeness of context, I_{CTX} , while allowing for different values depending on the trivia game information given to participants.

A noncompensatory model does not integrate context with individuating cues. Rather, either one or the other is the sole determiner of the PTJ. To give the noncompensatory model the best chance of succeeding, the PTJ was determined by the most diagnostic information: context or individuating cue information. If the diagnosticity of the cue is

equal to or greater than the informativeness of the context information, then the cue value alone is used to form the PTJ; otherwise, the context information alone is used. First, the diagnosticity of the cue (D_{CUE}) is calculated as the cue value's absolute different from .50, represented as T . Similarly, the diagnosticity of context information (D_{CTX}) is calculated as the absolute difference between the informativeness of context and T .

$$D_{CUE} = |T - V_{CUE}| \quad (3)$$

$$D_{CTX} = |T - I_{CTX}| \quad (4)$$

Knowing the diagnosticity of the cue value (D_{CUE}) and the context information (D_{CTX}), the noncompensatory model needs to arbitrate between them to determine which is the more diagnostic. This is performed by calculating the cue utility, U_{CUE} :

$$U_{CUE} = D_{CUE} - D_{CTX} \quad (5)$$

If U_{CUE} is greater than zero, then the diagnosticity of the cue is greater than the informativeness of the context, and so the cue value alone defines the PTJ. If it is less than zero, then the context informativeness is greater than the cue diagnosticity, and it alone defines the PTJ. Where cue and context diagnosticity were equal, individuating cue diagnosticity was arbitrarily chosen.

Only one issue remains: If U_{CUE} is less than zero and so context is used, then that value should depend on whether people are told that the trivia game was easy or difficult. Just as with the compensatory model, the value of context (CTX) is determined as per Equation (2) so that the value of CTX depends on the information given to participants.

Formally, the PTJ for the noncompensatory model is

$$PTJ \begin{cases} V_{CUE} & U_{CUE} \geq 0 \\ CTX & U_{CUE} < 0 \end{cases} \quad (6)$$

The two models, specified in Equations (1) and (6), are fitted to the aggregated data in Figure 3. The compensatory model yielded a parameter estimate of 59.83% for the context informativeness. The root-mean-square deviation (RMSD) for this model was 6.84% ($AIC = -1715$). RMSD is a measure of the difference between the behavioural data and the model fit. The noncompensatory model yielded a parameter estimate of 61.60% for the context informativeness. The RMSD for this model was 9.41% ($AIC = -1511$). There was no evidence of a statistically significant difference between the two models, $F(13, 13) = 1.89, p = .132$.²

²There are difficulties estimating the size of the effect when there are free parameters to be fitted (Roberts & Pashler, 2000), and so a standardized effect size is not offered.

Individual fits to the data

There is evidence that some people use compensatory strategies while others use noncompensatory strategies (e.g., Newell & Shanks, 2003). What is more, the poor fit of the noncompensatory model in the previous section may be because we are averaging across a set of step functions (as would be the case if people are noncompensatory), giving the appearance of a curve (as would be the case if people are compensatory). The compensatory and noncompensatory models were refitted to the data of each participant individually.

The mean parameter estimate for the compensatory model was 63.30% ($SD = 13.66$), whereas for the noncompensatory model, the parameter estimate was 65.72% ($SD = 12.19$). The RMSD for the compensatory model was 19.14% ($SD = 10.37, AIC = -1308$), and for the noncompensatory model, it was 18.75% ($SD = 9.77, AIC = -1310$); see Figure 4.

Figure 4c shows that the two models explain many participants almost equally (i.e., most data points fall on the diagonal), as suggested by a t -test comparison of the RMSD values, $t(79) = .25, p = .404$. Yet there are a small number of participants that are more likely to respond with a compensatory approach while others prefer a noncompensatory approach. We believe that understanding these individual differences in lie-truth judgment strategies will be an exciting area of future research.

So far, we have examined the ability of the two models to fit the data. At the aggregate level, the compensatory model performs better. At the individual level, which should be given more weighting in our inference, neither model appears preferable. But how generalizable are the two models? We use cross validation to assess this.

Eightfold cross validation

To test the performance and generalizability of the two models, we used eightfold cross validation so that 70 participants acted as the training set and 10 as the test set in each fold of the validation. The partitioning was performed randomly, but the same partitioning was used for assessing both models. In each fold, 500 randomly selected starting parameters bounded between .50 and 1.00 were used for parameter fitting to the training set. Additionally, the data partitions were generated in 500 different ways, at random. This yielded a parameter estimate of 61.63% ($SD = .72$) for the compensatory model and 61.23% ($SD = 1.17$) for the noncompensatory model, similar to the aggregate fits. The RMSDs of the trained models when trying to fit the smaller test sets were 26.19% ($SD = 3.73; AIC = -855.43$) and 27.77% ($SD = 3.63$), respectively.

A repeated measures t -test found evidence of a statistically significant difference between the RMSD values, $t(499) = -6.79, p < .001$. However, it should be borne in mind that both models do not generalize particularly well to new data sets, indicated by the relatively large RMSDs of around 27% error, and that the RMSD values are only just under two percentage points different from each other. Given these caveats, the

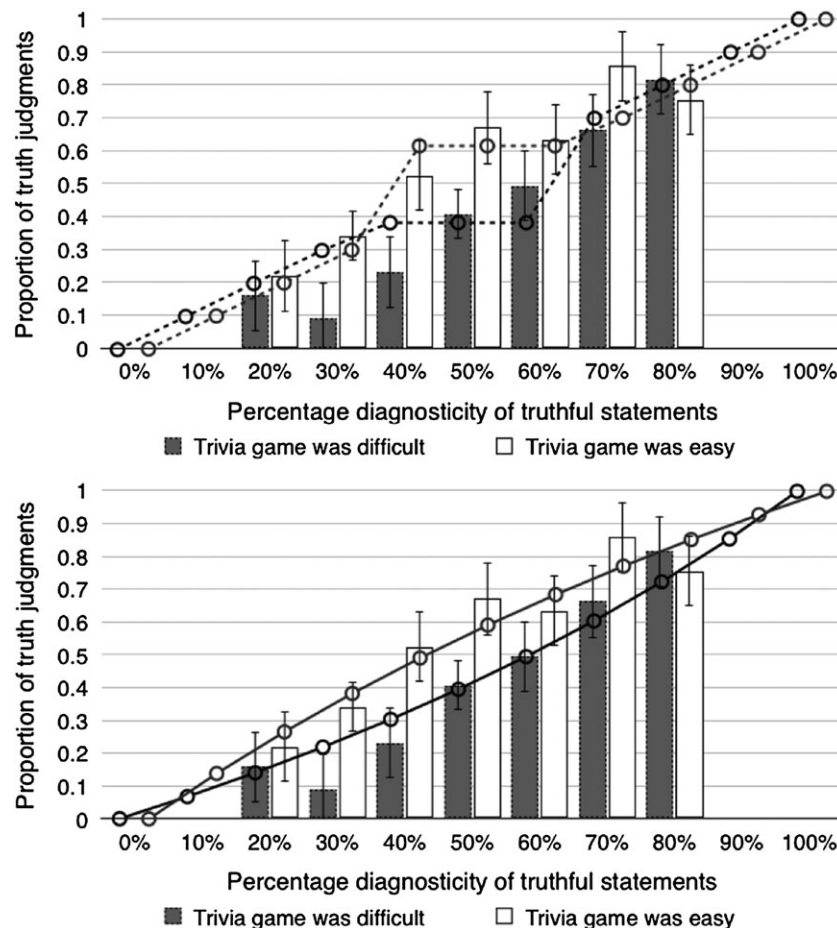


Figure 3. The model fits for the naïve Bayes compensatory (top) and noncompensatory models (bottom) depicted by lines. The models are overlaid onto the behavioral data and are depicted by columns. Error bars denote 95% confidence intervals of the behavioral data

compensatory model is the more generalizable of the two based on the individual data and predicts equally as well (at the individual level) as the noncompensatory model.

Summary

The aggregate model fits suggest that the compensatory model performs better than the noncompensatory model at the aggregate level, and the cross validation suggests that the compensatory model is the more generalizable to new data sets. The results from the individual model fits and the scatter plot in Figure 4c indicate that both models are performing similarly. We cautiously prefer a compensatory model—it predicts just as well as, if not better than, a noncompensatory model and generalizes somewhat more successfully. Although again it is worth remembering, the differences are small.

Regardless of which strategy is in play, if either, participants rely on their understanding of the current context when more individuating cues have low diagnosticity. Participants make use of context when the individuating cues have low diagnosticity, suggesting that context information is used as a means of making an informed guess when the available cues are weak.

DISCUSSION

Humans have a bad reputation for making good judgments. They are said to be inaccurate and misguided by their overwillingness to believe others (Bond & DePaulo, 2006; Levine et al., 1999; McCornack & Parks, 1986). A more optimistic view, and the one that this investigation put to the test, is that people make smart and adaptive judgments based on the information that is available.

Specifically, when the immediately available information (i.e., behavioral cues) does not support an accurate judgment, people shift toward using more generalized situation-specific context information to help guide their judgment. This shows up as a bias in the direction of the context. This is the prediction of the adaptive lie detector theory (Street, 2015; Street & Richardson, 2015). Consistent with the theory, we found that context had the least effect when the available cues were highly diagnostic of truth telling or lying and context had the greatest effect when the cues had low diagnosticity.

The findings dovetail with research in the social, decision making, and person perception literatures, which have demonstrated that people use the more individuating cues available to them when they have high diagnosticity but switch to other more general context-relevant knowledge when the individuating cues have low diagnosticity (Bar-Hillel, 1990;

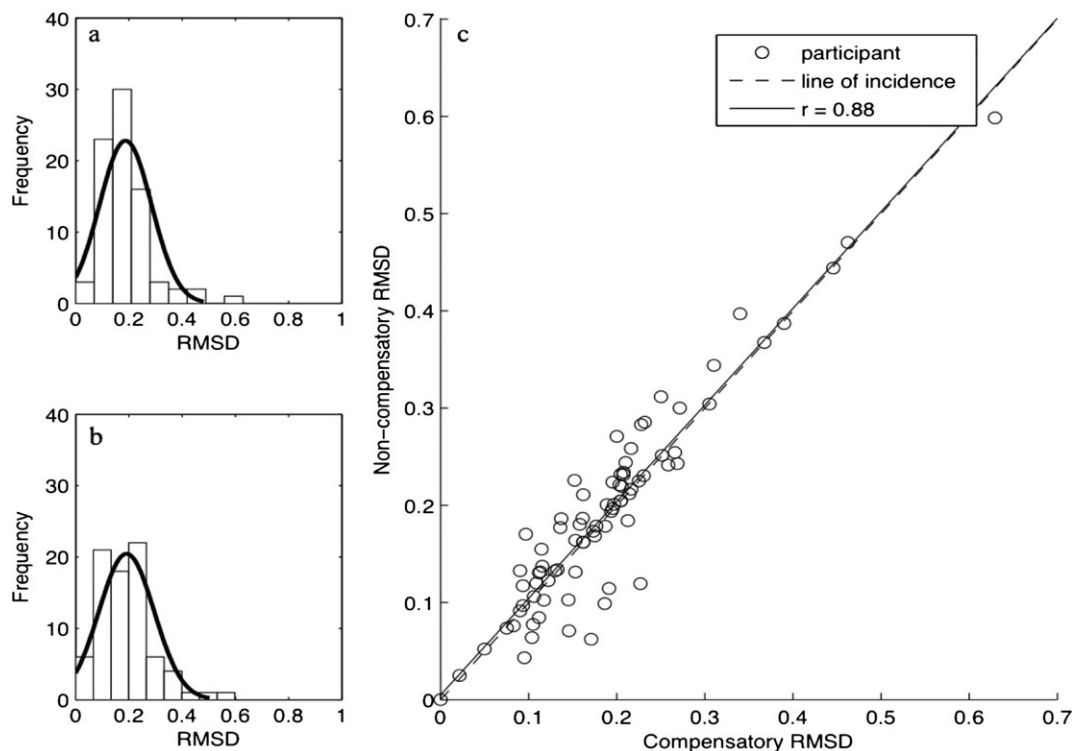


Figure 4. The distribution of errors (root-mean-square deviation values) when fitting individual participants to the (a) compensatory and (b) noncompensatory models. (c) The root-mean-square deviation for the compensatory and noncompensatory models for a given participant is plotted as a point. A point to the right of the dashed line of incidence indicates that the noncompensatory model provided a better fit than the compensatory model for that participant, and a point to the left of the line of incidence indicates the reverse

Garcia-Retamero & Rieskamp, 2008, 2009; Jekel et al., 2014; Kunda et al., 2002; Meiser et al., 2007). Judgments are considered to be adaptive in the sense that the information people use is dependent on the cue availability in the environment. This is a particularly Brunswikian approach inasmuch as the relative importance of one cue (context information) increases as the other cues decrease in diagnosticity (see Brunswik, 1952, on vicarious functioning). By shifting between different cues dependent on their diagnostic utility, people can make satisfactorily informed judgments, albeit with some errors (Simon, 1990).

It is perhaps also worth noting that our results are also consistent with a long tradition of research in category learning showing that attention shifts away from cues that either are not predictive of relevant outcomes or cause prediction errors (Kruschke, 2011; Le Pelley, Vadillo, & Luque, 2013).

Epistemic vigilance (Sperber et al., 2010; Sperber, 2013) has proposed a view similar to the adaptive lie detector proposal. To guard against potential deception, the epistemic vigilance account proposes that there is a module that uses more general person-based information such as the speaker's benevolence, their kindness, and so on (Mascaro & Sperber, 2009; Woolley & Ghossainy, 2013). Although not as abstracted as more general context information about all speakers, this person-based information is more generic than individuating cues because a speaker may appear benevolent in general, but this does not mean that they will always tell the truth. Put another way, benevolence can only tell one how likely the person is to lie in the long run but not whether the current statement they are producing is a lie or not. The

adaptive lie detector perspective aligns well with that of epistemic vigilance because it is usually the case that raters have to deal with low diagnostic cues (DePaulo et al., 2003), and so the use of more abstracted person-level information might allow for satisficing judgments (Glöckner, Hilbig, & Jekel, 2014; Simon, 1990) in the absence of more individuating cues. Note though that in our account when highly diagnostic information is available, raters do not rely on their "vigilance module" but rather base their judgment on the highly diagnostic information. That is, we do not argue for specialized mechanisms but rather argue for an importance of the interaction between the environment and the strategy (Simon, 1990).

We delved deeper into the lie-truth judgment strategy by considering two alternative strategies that could have explained our behavioral data. A naïve Bayes compensatory model assumed that people prefer to use both context and individuating cue information simultaneously but weighting them differently depending on which is the most diagnostic. A noncompensatory model assumed that people prefer to select either individuating cues or context alone to make their judgment. The models performed similarly, with slightly better performance and generalizability of the compensatory strategy of integrating context with cue information. Some previous work has found that people use noncompensatory strategies (Davis-Stober, Dana, & Budescu, 2010; Gaissmaier & Marewski, 2011; Gigerenzer, Hoffrage, & Kleinbölting, 1991; Gigerenzer, Martignon, Hoffrage, Rieskamp, & Czerlinski, 2008) but only when the task is structured so that they cannot freely choose how to use the available

information (Bröder, 2003; Newell & Shanks, 2003). There is quite some evidence supporting a compensatory approach to forming judgments (Bröder & Eichler, 2006; Newell & Fernandez, 2006; Newell & Shanks, 2004; Pohl, 2006; Richter & Späth, 2006; Söllner, Bröder, Glöckner, & Betsch, 2014), but there does seem to be contention about whether different individuals use compensatory or noncompensatory strategies (e.g., Newell & Shanks, 2003). The debate between whether people are compensatory or noncompensatory decision makers continues (Gigerenzer & Brighton, 2009; Hilbig & Richter, 2011; Pohl, 2011). We hope that other researchers will begin to engage with computational social cognition to address social perception and social reality (Korman, Voiklis, & Malle, 2015).

Although an adaptive lie detector position has been promoted here, it is important to note that, as we stated at the very beginning of this paper, adaptive judgments do not necessarily result in accurate judgments (Jussim, 2012). Of course, it is possible that the most diagnostic information available has only weak diagnostic value, and so accuracy will be low, despite selecting the best information available. Speakers can use persuasive techniques to mislead the listener (Cialdini, 2007; Evans, Barston, & Pollard, 1983), and salient information (Platzer & Bröder, 2012; Bond et al., 2013) or even the raters' own fluency biases (Schwarz, 2015) may mistakenly lead them astray.

Finally, we did not include a "no-context-information" condition where participants were told nothing about how likely people are to cheat on the task. In such a condition, we suspect that participants would rely on the individuating cues, but as the diagnosticity of these cues reduced, they would rely on their own interpretation of the context. Some people might be quite suspicious of trivia game players and so show a bias toward judging others as liars when the cue diagnosticity is low. But we suspect that people are more likely to interpret the situation such that people will not lie. After all, in our daily lives people rarely lie (DePaulo et al., 1996; Halevy et al., 2013). More importantly, studies over the years have consistently found that people are biased toward believing that others will tell the truth (Bond & DePaulo, 2006). Thus, we anticipate participants in such a no-context-information condition to respond similarly to the participants in our study who were told that the trivia game was easy, and so players would later be likely to tell the truth.

CONCLUSIONS

The adaptive lie detector position characterizes the response bias as a result of making an informed guess when unsure (Street, 2015; Street & Richardson, 2015). Thus, it is argued that researchers may not wish to remove biases from people's judgments, contrary to the bias-as-error position (Croskerry, Singhal, & Mamede, 2013; Gilbert, 1991; Larrick, 2004), because these biases are actually adaptive (Levine, 2014; Street & Richardson, 2014). It is interesting to note that attempts to remove biases from judgments by promoting rational thinking or by means of incentives often fail (Schwarz, Sanna,

Skurnik, & Yoon, 2007) because, as we would speculate, these biases are not errors but are functional and contribute to making an informed decision. Thus, if increasing lie detection accuracy is the goal, then efforts should be made not only to ensure that raters are making informed, logical judgments but also to ensure that the available cues are highly diagnostic (Vrij & Granhag, 2012).

In conclusion, the adaptive lie detector position tested and supported here proposes that people make smart, informed guesses in a low diagnostic world. When the immediately available individuating cues are highly diagnostic, people use those to inform their judgment. But as the cues decrease in diagnosticity, the role of more abstracted context-relevant knowledge contributes to the judgment, and this appears as a response bias. Thus, it is argued that response biases are not an error in judgment that must be suppressed but are signs of an adaptive strategy to make the most informed decision.

ACKNOWLEDGEMENTS

We would like to thank Arndt Bröder for his comments on the design of the study, Bruno Zumbo for his statistical guidance, and George Kachkovski and Sarah Maklad for their efforts in data collection. This work was supported by Federal Canadian grants to A. K. from the Natural Science and Engineering Research Council and the Social Sciences and Humanities Research Council.

REFERENCES

- Bar-Hillel, M. (1990). Back to base-rates. In R. M. Hogarth (Ed.), *Insights in decision making: A tribute to Hillel J Einhorn* (pp. 200–216). Chicago, IL: University of Chicago Press.
- Blair, J. P. (2006). From the field: Can detection of deception response bias be manipulated? *Journal of Crime and Justice*, 29(2), 141–152.
- Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review*, 10(3), 214–234.
- Bond, C. F., Howard, A. R., Hutchison, J. L., & Masip, J. (2013). Overlooking the obvious: Incentives to lie. *Basic and Applied Social Psychology*, 35, 212–221.
- Bröder, A. (2000). Assessing the empirical validity of the "take-the-best" heuristic as a model of human probabilistic inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(5), 1332–1346.
- Bröder, A. (2003). Decision making with the "adaptive toolbox": Influence of environmental structure, intelligence, and working memory load. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(4), 611–625.
- Bröder, A., & Eichler, A. (2006). The use of recognition information and additional cues in inferences from memory. *Acta Psychologica*, 121, 275–284.
- Brunswick, E. (1952). *The conceptual framework of psychology*. Chicago: University of Chicago Press.
- Burgoon, J. K., & Buller, D. B. (1994). Interpersonal deception: III. Effects of deceit on perceived communication and nonverbal behavior dynamics. *Journal of Nonverbal Behavior*, 18(2), 155–184.
- Cialdini, R. B. (2007). *Influence: The psychology of persuasion*. London: HarperCollins.
- Clark, H. H. (1996). *Using language*. New York: Cambridge University Press.

- Clark, H. H., Schreuder, R., & Buttrick, S. (1983). Common ground and the understanding of demonstrative reference. *Journal of Verbal Learning and Verbal Behavior*, 22, 245–258.
- Croskerry, P., Singhal, G., & Mamede, S. (2013). Cognitive debiasing 1: Origins of bias and theory of debiasing. *British Medical Journal Quality and Safety*, 22, 58–64.
- D'Errico, J. (2012). fminsearchbnd, fminsearchcon – file exchange – MATLAB central [online forum content]. Retrieved from <http://www.mathworks.ca/matlabcentral/fileexchange/8277-fminsearchbnd-fminsearchcon>.
- Davis-Stober, C. P., Dana, J., & Budescu, D. V. (2010). Why recognition is rational: Optimality results on single-variable decision rules. *Judgment and Decision Making*, 5(4), 216–229.
- Degani, A., Shafto, M., & Kirlik, A. (2006). What makes vicarious functioning work? Exploring the geometry of human-technology interaction. In A. Kirlik (Ed.), *Adaptive perspectives on human-technology interaction* (pp. 179–196). New York: Oxford University Press.
- DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology*, 70(5), 979–996.
- 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.
- DePaulo, P. J., & DePaulo, B. M. (1989). Can deception by salespersons and customers be detected through nonverbal behavioral cues? *Journal of Applied Social Psychology*, 19(18), 1552–1577.
- Evans, J. S. B. T., Barston, J. L., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory and Cognition*, 11, 295–306.
- Fiedler, K., Armbruster, T., Nickel, S., Walther, E., & Asbeck, J. (1996). Constructive biases in social judgment: Experiments on the self-verification of question contents. *Journal of Personality and Social Psychology*, 71(5), 861–873.
- Gaissmaier, W., & Marewski, J. N. (2011). Forecasting elections with mere recognition from small, lousy samples: A comparison of collective recognition, wisdom of crowds, and representative polls. *Judgment and Decision Making*, 6(1), 73–88.
- Garcia-Retamero, R., & Rieskamp, J. (2008). Adaptive mechanisms for treating missing information: A simulation study. *The Psychological Record*, 58, 547–568.
- Garcia-Retamero, R., & Rieskamp, J. (2009). Do people treat missing information adaptively when making inferences? *The Quarterly Journal of Experimental Psychology*, 62, 1991–2013.
- Gigerenzer, G., & Brighton, H. (2009). Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science*, 1(1), 107–143.
- Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, 103(4), 650–669.
- Gigerenzer, G., Hertwig, R., & Pachur, T. (2011). *Heuristics: The foundations of adaptive behavior*. Oxford: Oxford University Press.
- Gigerenzer, G., Hoffrage, U., & Kleinbölting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, 98, 506–528.
- Gigerenzer, G., Martignon, L., Hoffrage, U., Rieskamp, J., & Czerlinski, J. (2008). One-reason decision making. In C. R. Plott & V. L. Smith (Eds.), *Handbook of experimental economics results*, 1 (pp. 1004–1017). Oxford: Elsevier.
- Gilbert, D. T. (1991). How mental systems believe. *American Psychologist*, 46(2), 107–119.
- Ginossar, Z., & Trope, Y. (1980). The effects of base rates and individuating information on judgments about another person. *Journal of Experimental Social Psychology*, 16, 228–242.
- Glöckner, A., Hilbig, B. E., & Jekel, M. (2014). What is adaptive about decision making? A parallel constraint satisfaction account. *Cognition*, 133, 641–666.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics* (Vol. 3, pp. 41–58). New York: Academic Press.
- Halevy, R., Shalvi, S., & Verschuere, B. (2013). Being honest about dishonesty: Correlating self-reports and actual lying. *Human Communication Research*, 40(1), 54–72.
- Heyes, C. (2014). Submentalizing: I am not really reading your mind. *Perspectives on Psychological Science*, 9(2), 131–143.
- Hilbig, B. E., & Richter, T. (2011). Homo heuristicus outnumbered: Comment on Gigerenzer & Brighton (2009). *Topic in Cognitive Psychology*, 3(1), 187–196.
- Jekel, M., Glöckner, A., Bröder, A., & Maydych, V. (2014). Approximating rationality under incomplete information: Adaptive inferences for missing cue values based on cue-discrimination. *Judgment and Decision Making*, 9(2), 129–147.
- Jussim, L. (2012). *Social perception and social reality: Why accuracy dominates bias and self-fulfilling prophecy*. Oxford: Oxford University Press.
- Koehler, J. J., (1996). The base rate fallacy reconsidered: Descriptive, normative, and methodological challenges. *Behavioral and Brain Sciences*, 19, 1–53.
- Korman, J., Voiklis, J., & Malle, B. F. (2015). The social life of cognition. *Cognition*, 135, 30–35.
- Kruschke, J. K. (2011). Models of attentional learning. In E. M. Pothos & A. J. Wills (Eds.), *Formal approaches in categorization* (pp. 120–152). Cambridge, UK: Cambridge University Press.
- Kunda, Z., Davies, P. G., Adams, B. D., & Spencer, S. J. (2002). The dynamic time course of stereotype activation: Activation, dissipation, and resurrection. *Journal of Personality and Social Psychology*, 82(3), 283–299.
- Larrick, R. P. (2004). Debiasing. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 316–337). Oxford: Blackwell Publishing.
- Le Pelley, M. E., Vadillo, M. A., & Luque, D. (2013). Learned predictiveness influences rapid attentional capture: Evidence from the dot probe task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39, 1888–1900.
- Levine, T. R. (2014). Truth-default theory (TDT): A theory of human deception and deception detection. *Journal of Language and Social Psychology*, 33, 378–392.
- Levine, T. R., Park, H. S., & McCormack, S. A. (1999). Accuracy in detecting truths and lies: Documenting the “veracity effect”. *Communication Monographs*, 66, 125–144.
- Mandelbaum, E. (2014). Thinking is believing. *Inquiry: An interdisciplinary Journal of Philosophy*, 57(1), 5–96.
- Mascaro, O., & Sperber, D. (2009). The moral, epistemic, and mindreading components of children’s vigilance towards deception. *Cognition*, 112, 367–380.
- Masip, J., Alonso, H., Garrido, E., & Herrero, C. (2009). Training to detect what? The biasing effects of training on veracity judgments. *Applied Cognitive Psychology*, 23, 1282–1296.
- McCormack, S. A., & Parks, M. R. (1986). Deception detection and relationship development: The other side of trust. In M. L. McLaughlin (Ed.), *Communication yearbook 9* (pp. 377–389). Beverly Hills: Sage.
- Meiser, T., Sattler, C., & von Hecker, U. (2007). Metacognitive inferences in source memory judgements: The role of perceived differences in item recognition. *The Quarterly Journal of Experimental Psychology*, 60(7), 1015–1040.
- GMeissner, C. A., & Kassir, S. M. (2002). “He’s guilty!”: Investigator bias in judgments of truth and deception. *Law and Human Behavior*, 26(5), 469–480.
- Millar, G. M., & Millar, K. U. (1997). The effects of cognitive capacity and suspicion on truth bias. *Communication Research*, 24(5), 556–570.
- Moston, S., Stephenson, G. M., & Williamson, T. M. (1992). The effects of case characteristics on suspect behaviour during police questioning. *British Journal of Criminology*, 32(1), 23–40.

- Nadarevic, L., & Erdfelder, E. (2013). Spinoza's error: Memory for truth and falsity. *Memory & Cognition*, 41, 176–186.
- Nakagawa, S., & Schielzeth, H. (2010). Repeatability for Gaussian and non-Gaussian data: A practical guide for biologists. *Biological Reviews*, 85, 935–956.
- Newell, B. R., & Fernandez, D. (2006). On the binary quality of recognition and the inconsequentiality of further knowledge: Two critical tests of the recognition heuristic. *Behavioral Decision Making*, 19(4), 333–346.
- Newell, B. R., & Shanks, D. R. (2003). Take the best of look at the rest? Factors influencing “one-reason” decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(1), 53–65.
- Newell, B. R., & Shanks, D. R. (2004). On the role of recognition in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 923–935.
- O'Sullivan, M. (2003). The fundamental attribution error in detecting deception: The boy-who-cried-wolf effect. *Personality and Social Psychology Bulletin*, 29, 1316–1327.
- Platzer, C., & Bröder, A. (2012). Most people do not ignore salient invalid cues in memory-based decisions. *Psychonomic Bulletin and Review*, 19, 654–661.
- Pohl, R. F. (2006). Empirical tests of the recognition heuristic. *Journal of Behavioral Decision Making*, 19, 251–271.
- Pohl, R. F. (2011). On the use of recognition in inferential decision making: An overview of the debate. *Judgment and Decision Making*, 6(5), 423–438.
- Richter, T., & Späth, P. (2006). Recognition is used as one cue among others in judgment and decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32, 150–162.
- Roberts, S., & Pashler, H. (2000). How persuasive is a good fit? A comment on theory testing. *Psychological Review*, 107(2), 358–367.
- Schwarz, N. (2015). Metacognition. In M. Mikulincer, P. R. Shaver, E. Borgida & J. A. Bargh (Eds.), *APA handbook of personality and social psychology: Attitudes and social cognition* (pp. 203–229). Washington, DC: APA.
- Schwarz, N., Sanna, L. J., Skurnik, I., & Yoon, C. (2007). Metacognitive experiences and the intricacies of setting people straight: Implications for debiasing and public information campaigns. *Advances in Experimental Social Psychology*, 39, 127–161.
- Simon, H. A. (1990). Invariants of human behaviour. *Annual Review of Psychology*, 41, 1–20.
- Söllner, A., Bröder, A., Glöckner, A., & Betsch, T. (2014). Single-process versus multiple-strategy models of decision making: Evidence from an information intrusion paradigm. *Acta Psychologica*, 146, 84–96.
- Sperber, D. (2013). Speakers are honest because hearers are vigilant: Reply to Kourken Michaelian. *Episteme*, 10(1), 61–71.
- Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origg, G., & Wilson, D. (2010). Epistemic vigilance. *Mind & Language*, 25(4), 359–393.
- Sporer, S. L., & Schwandt, B. (2006). Paraverbal indicators of deception: A meta-analytic synthesis. *Applied Cognitive Psychology*, 20, 421–446.
- Sporer, S. L., & Schwandt, B. (2007). Moderators of nonverbal indicators of deception: A meta-analytic synthesis. *Psychology, Public Policy, and Law*, 13(1), 1–34.
- Street, C. N. H. (2015). ALIED: Humans as adaptive lie detectors. *Journal of Applied Research in Memory and Cognition*. Advance online publication. doi: 10.1016/j.jarmac.2015.06.002.
- Street, C. N. H., & Richardson, D. C. (2014). Lies, damn lies, and expectations: How beliefs about the base rates inform lie-truth judgments. *Applied Cognitive Psychology*, 29(1), 149–155.
- Street, C. N. H., & Richardson, D. C. (2015). Descartes versus Spinoza: Truth, uncertainty, and bias. *Social Cognition*, 33(3), 227–239.
- Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 103–130). New Jersey: Lawrence Erlbaum.
- Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities* (2nd ed.). New Chichester: John Wiley.
- 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*, 2(1), 110–117.
- Woolley, J. D., & Ghossainy, M. E. (2013). Revisiting the fantasy-reality distinction: Children as naïve skeptics. *Child Development*, 84(5), 1496–1510.
- Yzerbyt, V. Y., Schadron, G., Leyens, J.-P., & Rocher, S. (1994). Social judgeability: The impact of meta-informational cues on the use of stereotypes. *Journal of Personality and Social Psychology*, 66(1), 48–55.

Authors' biographies:

Chris N. H. Street obtained his PhD from University College London. He is a Lecturer at the University of Huddersfield and is a Member of the International Research Centre for Investigative Psychology. His research explores the cognitive processes underlying dishonesty and situations of social conflict. His Twitter handle is @cnhstreet.

Walter F. Bischof obtained his PhD from the University of Bern, Switzerland. He is a Professor Emeritus at the University of Alberta in Edmonton, Canada, and an Adjunct Professor at the University of British Columbia, Canada. His research bridges computer science and psychology, with particular interests in human–computer interaction, visual attention, eye movements, and visual perception.

Miguel A. Vadillo obtained his PhD from the University of Deusto (Bilbao, Spain). He is a Lecturer at King's College London in the UK and Honorary Research Fellow at University College London. His research considers aspects of learning, memory, and decision making. He currently blogs at mvadillo.com/de-mente/.

Alan Kingstone obtained his PhD from the University of Manchester, UK. He is a Professor at the University of British Columbia in Vancouver, Canada, and Director of the Brain and Attention Research Lab. His research takes an ethological approach to cognition, spanning the areas of social cognition and visual attention.

Authors' addresses:

Chris N. H. Street, University of Huddersfield, UK.

Walter F. Bischof, University of Alberta, Canada.

Miguel A. Vadillo, King's College London, UK.

Alan Kingstone, University of British Columbia, Canada.