Aligning Spinoza with Descartes:

An Informed Cartesian Account of the Truth Bias

Chris N. H. Street[1,2], & Alan Kingstone[3]

[1]Department of Psychology, University of Huddersfield, Huddersfield, UK, HD1 3DH

[2]International Research Centre for Investigative Psychology, University of Huddersfield, Huddersfield, UK, HD1 3DH

[3]Department of Psychology, University of British Columbia, Vancouver BC, Canada, V6T 1Z4

Correspondence should be address to:

Chris N. H. Street

Department of Psychology

University of Huddersfield,

Huddersfield, UK

HD1 3DH.

Email: c.street@hud.ac.uk

Telephone: +44 1484 472631

Abstract

There is a bias towards believing information is true rather than false. The Spinozan account claims there is an early, automatic bias towards believing. Only afterwards can they engage in an effortful re-evaluation and disbelieve the information. Supporting this account, there is a greater bias towards believing information is true when under cognitive load. However, developing on the Adaptive Lie Detector (ALIED) theory, the informed Cartesian can equally explain this data. The account claims the bias under load is not evidence of automatic belief; rather, people are undecided, but if forced to guess they can rely on context information to make an informed judgment. The account predicts, and we found, that if people are not forced to guess when under load, but rather can explicitly indicate their uncertainty, there should be no bias towards believing because they are no longer required to guess. Thus we conclude that belief formation can be better explained by an informed Cartesian account.

Keywords: Adaptive Lie Detector; Informed Cartesian; Lie detection; Spinozan mind; Truth bias; Uncertainty.

How do people decide what to assent belief in and what to dismiss as unbelievable? According to the 'Spinozan' account (Gilbert, Krull, & Malone, 1990), people have no choice but to first believe in everything, and then, only afterwards, can they decide whether or not to disbelieve it (Gilbert et al., 1990; see also Levine, 2014; for a weaker defaulting claim that is consistent with the Spinozan account). This account has been influential in a variety of research areas, from psycholinguistics to lie detection (see Gilbert, 1991; and Mandelbaum, 2014, for an overview). An alternative possibility, known as the 'Cartesian' account, proposes that there is no automatic need to first believe in something. Instead, people can hold information in mind without attaching any belief value to it, and then assess the credibility of that information as believable or not. Under this account, there is an initial period of uncertainty, rather than an initial automatic certainty of belief. Recently, there has been an informal debate in the literature with some citing evidence in favour and others citing evidence against the Spinozan account. This article aligns the apparently contradictory findings under a unified 'informed Cartesian' account.

In their original set of experiments, Gilbert et al. (1990) showed participants videos of real or fake smiles. After each trial during this learning phase, participants were told explicitly whether the expression was honest or deceptive. On some trials, participants performed a concurrent task: To respond as quickly as possible to a tone (Study 1) or identify a clearly higher versus lower pitched tone (Study 2), both of which were sufficient to act as a cognitive load.

The Spinozan view holds that when the facial expressions were first encountered, they were automatically believed to be truthful, and it was only during a second active and effortful stage that the credibility of that initial belief was assessed. Within this Spinozan framework, the following prediction was made. First, the researchers reasoned that an increase in cognitive load (arising from the concurrent probe tone task) would result in the effortful stage being bypassed, and as a result, judgements would reflect only the initial belief-assenting stage (Gilbert et al., 1990; Mandelbaum, 2014). Consistent with this prediction, the results revealed that during a subsequent test phase, when participants were given a true or false two-alternative forced-choice task, they were more likely to believe emotional expressions were genuine if they had to cope with a secondary

task during the learning phase (Gilbert, 1991; Gilbert et al., 1990; Mandelbaum, 2014; cf. Hasson, Simmons, & Todorov, 2005).

Because the automatic belief stage must precede the more evaluative stage, the Spinozan account also predicts that the truth bias should appear early in the judgement process: Over time, the bias should decline because of a shift towards more evaluative processing (Gilbert, 1991; Gilbert, Tafarodi & Malone, 1993; Masip, Garrido & Herrero, 2006; cf. Skurnik, Yoon, Park & Schwarz, 2005). The account relies heavily on reaction time data from psycholinguistic research (Mandelbaum, 2014). Indeed, Gilbert et al. (1990, Study 2; 1993, Study 3) have put this temporal prediction of theirs to the test. The data supported this prediction, with the truth bias being present when under time pressure, but waned without time pressure (Gilbert et al., 1993). They also found that forewarning participants of false information was not sufficient to prevent an initial truth belief (Gilbert et al., 1990).

A recent study also put this temporal prediction to the test (Street & Richardson, 2015a). They found a truth bias early in the decision process, as predicted by a Spinozan account. In their study, participants were forced to make a response at the onset of the stimulus and to continue updating their button response over time, depending on whether they changed their mind. This gave a temporal profile of the decision process. On the basis of the Spinozan account, the authors predicted a truth bias to be observed early in the judgement, resulting from an automatic and early belief. As predicted, a truth bias was observed during the initial moments of decision-making. But they also found that this early truth bias disappeared when participants were provided with the additional option to respond that they were 'unsure' about the credibility of the information. The Spinozan prediction was that the inclusion of this third option would be irrelevant; that is because is an automatic and necessary early bias towards believing. Street and Richardson proposed that an 'informed Cartesian' account readily explains this new finding, and the previous data as well. Specifically, the informed Cartesian view is that people make use of context information (e.g., 'people usually tell the truth'; see DePaulo, Kashy, Kirkendol, Wyer, & Epstein, 1996; Grice, 1975; McNally & Jackson, 2013) to make the best guess. 'Context' refers to information that pertains to this situation in general, giving an indication of whether most people will lie or tell the truth. When

participants are unsure but are required to make a response – 'True or False' – that context information is expressed as a truth bias. Accordingly, when a third 'Unsure' response option is provided, there is no longer a need to make an informed guess – people can instead indicate their uncertainty. The 'True' response shifts to 'Unsure' and with that, the early truth bias disappears. In other words, what had appeared to be an automatic early bias to believe everything is true (the Spinozan account) is revealed to be a context-driven best guess (the informed Cartesian account). The informed Cartesian account is consonant with the Adaptive Lie Detector (ALIED) theory (Street, 2015), which proposes that the bias towards believing results from an attempt to make a context-informed guess when uncertain.

The informed Cartesian account proposes that interrupting the encoding process during learning (e.g., by having participants respond to a noise while simultaneously learning whether information is true or false) takes resources or attention away from the learning task and causes uncertainty. This account adopts a boundedly rational approach (Simon, 1990). A boundedly rational agent is one who makes satisfactorily accurate responses with limited information and limited cognitive resources. Under uncertainty, people use heuristics (e.g., 'rules of thumb') to make sense of the world (Gigerenzer, Todd, & The ABC Research Team, 1999). For instance, when people are lacking information about a particular individual, they rely on more generalized stereotypic knowledge to make an informed impression of the person (Lick & Johnson, 2014). In the case of lie detection, one may rely on a more generalized understanding of how likely people are to lie in a given context, and use that to inform a judgement when one is unsure about whether any given person is lying (see Street, 2015, for a review). Because most people tell the truth most of the time (DePaulo et al., 1996), using this generalized heuristic will lead to a bias towards believing people are telling the truth. Put simply, if participants are unsure about something but are forced to guess, the informed Cartesian claims that they do not flip a mental coin – rather, they rely on their understanding of the context to make their guess.

From this it also follows that it need not always be the case that people will expect others to be telling the truth, but that the expectation will vary with the context. For instance, police officers tend to be biased towards judging suspects' statements to be lies

(Meissner & Kassin, 2002). The informed Cartesian account readily explains this as a context-dependent heuristic: A belief that suspects in general tend to lie (Moston, Stephenson, & Williamson, 1992). Note that while the effect of context is irrelevant to the Spinozan account, it is foundational to an informed Cartesian account. Decisions are derived from prior experience, knowledge, and an understanding of the current situation (Deutsch, Kordts-Freudinger, Gawronski, & Strack, 2009; Hasson et al., 2005; Richter, Schroeder, & Wöhrmann, 2009; Schroeder, Richter, & Hoever, 2008). This point is supported by the finding that if one *changes the context* so that raters believe most people will lie, there is an early bias towards *dis*believing the speaker (i.e., a change in bias; Mayo, Schul, & Burnstein, 2004; Richter et al., 2009; Street & Richardson, 2015b; see also Levine, Kim, & Blair, 2010; Sperber et al., 2010). According to the Spinozan framework, this type of systematic movement in the judgement should not occur – initially people should always be biased to believe, regardless of context. According to the informed Cartesian account, there will only be an initial bias to believe when the context suggests that most people will tell the truth and they are then forced to make a decision, as shown by Street and Richardson (2015a, b).

One might, however, seek to account for these new data by modifying the Spinozan position, and arguing that in all the above situations participants were not under a dual-task cognitive load. A variant of the Spinozan position might propose that the Spinozan account predicts that *only* cognitive load sufficiently taxes the conscious mind to allow unconscious processes to be seen, and that the account makes no time-based predictions.[1] The early stages of processing may have reflected the involvement of a more cognitive controlled second stage, and this tended to 'bleed' into the early moments of the decision.

Now, strictly speaking, this bleeding should not happen according to the Spinozan account (Gilbert, 1991). The original account claims an automatic (in the sense of fast, unintentional, and uncontrollable; Mandelbaum, 2014) belief stage *precedes* an evaluation stage (Gilbert et al., 1990, 1993; Skurnik et al., 2005; see also Unkelbach, 2007). Gilbert (1991) notes that the evaluative and automatic belief stages are modular and so are strictly independent. Nonetheless, the most straightforward and scientific way

[1] We are unaware of any published claim of this sort. However, an anonymous reviewer of a previous manuscript suggested just such a modified Spinozan account.

to test this bleeding Spinozan account would be to simply put it to the test and give it every chance to succeed. Thus, the present study was adapted from the method employed by Gilbert et al. (1990, Study 2). In their study, they created a cognitive load on some trials by having participants perform a dual task. They found that people were biased towards believing when there was a cognitive load, which they took as supporting the Spinozan account. We used this same method here. Importantly, we added a between-subjects manipulation. Participants were either given a two-alternative forced lie-truth choice or were given an additional response option of saying they were unsure.

In line with Street and Richardson (2015a), we predicted that there would be a bias towards believing statements are true on interrupted items and when forced to make a judgement because they would rely on context-relevant information to help make the best guess when unsure. But when able to indicate their uncertainty, and no longer have to make the guess, we predicted no truth bias would be observed. Critically, if the original or modified Spinozan account is correct, the uncertainty response option should only eliminate the truth bias when participants are given the cognitive load task. When they are under a cognitive load, the evaluation stage is bypassed, and therefore, people must hold a truth belief. As a result, the presence of a third uncertainty response option will be irrelevant and the truth bias will be observed. Following Street and Richardson (2015a), the unsure responses are removed from analysis because we are interested in whether people made more truth than lie judgements, regardless of their unsure responses.

Methods

Participants

Eighty-two undergraduate students (60 female) were recruited from a Canadian university (M_{age} = 21.35, SD = 3.92, range 17-44 years) and were paid $5.

Materials

We used the Bloomsbury Deception Set (Street et al., 2011) where each speaker delivered one honest and one deceptive statement about people they had supposedly met on holiday, under the guise that the researchers were filming for a documentary. The two statements from each speaker were organized into two separate stimulus sets so that each speaker either told the truth or lied in each stimulus set. There were 10 lies and 10 truths

in each set, and each participant only viewed one set. Full details of the stimuli can be found in Street and Richardson (2015b). The stimuli are available by request.

Only the first 8 s of each clip was used in the current study, mirroring the 8-s presentation time of Gilbert et al. (1990). The audio was removed from each clip, leaving participants to judge deception from non-verbal behaviour alone. This is in keeping with the emotional expression experiment of Gilbert et al. (1990) where participants were shown a brief non-normed video of a smile without audio. The order of the clips was randomized once before the first run of the experiment and that order was maintained for every participant, following Gilbert et al.'s (1990) procedure.

The videos were used during a training phase. In Gilbert et al.'s (1990) Study 2, participants were shown the same video a second time and asked to judge if the video was deceptive or honest. But because presenting the video a second time may lead to a re-evaluation of whether the speaker was telling the truth or not, to give the modified Spinozan account every chance to succeed, in the test phase we took still images from the video clip. Participants were told that they should use the image to recall back to the video and remember whether it was a truth or a lie. To further prevent the possibility that they would attempt to re-evaluate their judgements based on potential 'tells' to deception in the image (e.g., lack of eye contact), each video clip was reviewed manually and the still frame of the person found to be in as neutral a pose as could be found was used as the stimulus image in the test phase; for example, their arms were not gesticulating where possible and their face was relatively expressionless (Birmingham, Bischof, & Kingstone, 2008, 2009).

Procedure

The procedure was adapted from Gilbert et al., (1990). During a learning phase, participants saw an 8-s clip of a speaker. Then, the word 'Lie' or 'Truth' was shown on screen for 3 s (referred to here as the feedback). There were a total of 20 trials, half of which were true and half false. On eight of the 20 trials there was a 500-Hz tone that sounded after 750 ms of the appearance of the feedback word, and lasted 390 ms[2].

[2] The Spinozan account has never addressed the time scale at which belief formation occurs; only that it occurs automatically and rapidly. The timings used here were taken from Gilbert et al.'s (1990) studies supporting the Spinozan account.

Participants were instructed to hit the spacebar as quickly as possible when they heard the tone, ostensibly because we were interested in how the speed of response relates to memory (as in Gilbert et al., 1990; Study 1). In reality, the tone was intended to increase the cognitive load and interrupt processing of the feedback word. Note also that Gilbert et al. (1990) found this tone task was sufficient to engender a cognitive load. The first two and last two trials were never accompanied with a tone, and were not presented at the test phase to reduce primacy and recency effects. Of the remaining 16 trials, half of those were accompanied by the tone. They eight trials chosen to be accompanied by the tone were selected at random for each new participant.

During a test phase, participants would have to identify whether each speaker had been lying or telling the truth. On each trial, a still image of a speaker was shown. Importantly, participants were instructed they should not attempt to detect deception from the image provided because the image had been taken from a different statement and so contained no useful information. They were told that the photograph was there just to indicate which speaker should be recalled. They had to think back to the video phase of the experiment and recall whether the feedback indicated if the person had lied or told the truth.

Our experiment introduced a between-subjects manipulation into the original design. Half of our participants were randomly assigned to the lie-truth forced-choice condition (LT condition; $n = 42$): they were to judge whether the speaker lied or told the truth. The other half were assigned to the lie-truth-unsure unforced choice condition (LTU condition; $n = 39$): they could judge the speaker as having lied, told the truth, or respond that they were unsure if the speaker had lied or told the truth.

Design

The independent variables were the forced-choice response condition (LT or LTU, between participants), the veracity of the statement (truth or lie, within subjects), and the dual-task tone interruption, which served as the cognitive load manipulation (cognitive load or no load, within subjects).

The number of truth judgements made out of all the lie-truth judgements was our dependent variable, which is referred to as the proportion of truth judgements (PTJ). Note that this calculation for the LTU condition discards all the unsure responses. This means

that a PTJ of 0.50 represents a lack of bias both the LT and LTU conditions, and values higher than this in both conditions indicate a bias towards making truth over lie judgements. 10% of responses were unsure responses. By ignoring the unsure responses, the two conditions could be directly compared.

Results

As expected by a Spinozan account, raters were truth-biased when judging items that were interrupted during learning, but not when judging uninterrupted items. This Spinozan effect was only found when raters made a forced binary lie-truth judgement. When they could indicate that they were unsure, interruption did not result in a greater proportion of truth to lie judgements. This invalidates the Spinozan account, and the modified Spinozan account. Our results are best accommodated by the informed Cartesian account, which predicts Spinozan-like biases under specific conditions (e.g., when the response options do not permit individuals to explicitly state their uncertainty), in which case they make the best guess when the context information is factored in.

A 2 (response condition: LT or LTU, between subjects) x 2 (speaker veracity: lie or truth, within subjects) x 2 (cognitive load: load or no load during learning, within subjects) mixed ANOVA was conducted on the PTJ. The PTJ was calculated as the number of truth judgements divided by the total number of lie and truth judgements – after removing the unsure responses. Thus in both the LT and LTU conditions, a proportion of 0.50 truth judgements represents a lack of bias and values greater than 0.50 indicate a bias towards making truth judgements.

There was a main effect of veracity, $F(1, 79) = 55.07$, $p < .001$, $\eta_p^2 = 0.41$: People made more truth judgments when rating truthful ($M = 0.64$, $SD = 0.21$, 95% CI [0.60, 0.69]) rather than deceptive statements ($M = 0.37$, $SD = 0.22$, [0.33, 0.42]). That is, people remembered who lied and told the truth above chance rates.

Most critical to our hypothesis was that the LTU condition would not show an increased truth bias under cognitive load because they are not forced to guess when unsure.[3] A response condition x cognitive load interaction confirmed this prediction, $F(1,$

[3] We should note that one might wish to suggest that because participants experienced half truthful and half deceptive items in the learning phase, they may have anticipated an equal base rate in the test phase also, which could have given rise to the

79) = 5.77, p = .019, η_p^2 = 0.07. Bonferroni-corrected t-tests showed that raters in the LT condition were more likely to judge items as truthful when learning under cognitive load (M = 0.56, SD = 0.17, 95% CI [0.51, 0.61]) than when there was no load (M = 0.47, SD = 0.19, [0.41, 0.53]), with a medium effect size: t(41) = 2.56, p = .011, d = 0.50. So when forced to judge, people appear Spinozan. But in the LTU condition, there was no evidence of significantly more truth judgements when there was no load (M = 0.48, SD = .17, [0.43, 0.54]) compared to when there was a cognitive load, (M = 0.51, SD = 0.19, [0.45, 0.57]), t(38) = -0.84, p = .404, d = -0.17, Figure 1.[4] So when people are not forced to judge, they no longer look Spinozan. It would seem people appear Spinozan only when forced to judge. But when able to indicate their uncertainty people no longer show the biased responding. The analyses for all other main effects and interactions had a negligible effect size. They are not critical to the current hypotheses, but are reported in Appendix B.

Having the unsure response available reduced the proportion of truth responses, but not the proportion of lie responses. This is consistent with the informed Cartesian hypothesis that the relatively high PTJ is the result of being forced to make a guess when unsure.

Because we were interested in whether people were biased to make truth judgements, our dependent variable was the PTJ. However, Gilbert and colleagues used accuracy as their dependent variable, which offers a less direct assessment of bias. In the online supplement, we reanalyse the data using Gilbert et al.'s procedure. The LT condition closely replicates what Gilbert et al. found in their experiment.

In summary, people are biased towards believing information is true when interrupted (LT condition), as is traditionally expected of a Spinozan system – in as much

---

lack of biased responding in the unforced LTU condition. Critically, this proposal also predicts that no bias will be observed in the forced choice LT condition, a prediction that our data soundly disconfirm, and with it, the concern.

[4] A Bayes factor analysis was also conducted because it can give an insight into whether null effects are truly inconclusive (e.g., resulting from a lack of power) or whether they actually give confirmatory support for the lack of an effect. The analysis supports the lack of an effect, and is presented in Appendix A.

as it is considered to be a by-product of the architecture. An informed Cartesian should also show this bias, for a different reason – it claims that the bias reflects an informed guess when unsure. When no longer forced into judgement and able to indicate their uncertainty (LTU condition), people were no more likely to believe statements were true under cognitive load versus when there was no load. The findings cannot be explained by a Spinozan account, but are consistent with the predictions of an informed Cartesian account.

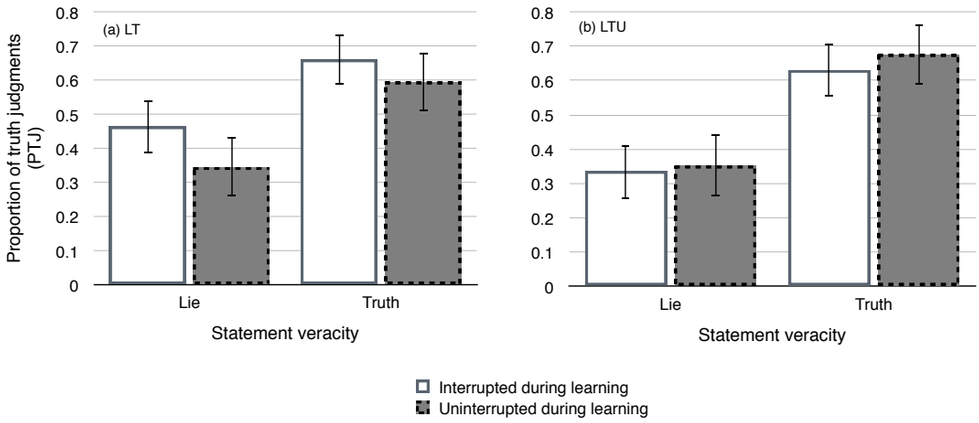

Figure 1. (a, left) The proportion of truth judgments (PTJ) made by raters in the LT forced choice condition. Raters in this condition showed a bias towards believing on items that were interrupted (white bars) compared to uninterrupted (grey bars) during learning. (b, right) The PTJ made by raters in the LTU unforced choice condition. Raters were no more biased towards believing on interrupted compared to uninterrupted items. Error bars denote 95% confidence intervals.

Discussion

Believing what others say is the unwavering default, according to the Spinozan position (Gilbert, 1991; Mandelbaum, 2014; see also Levine, 2014, for a weaker defaulting account). Only after assenting belief can people more critically consider whether they will continue believing or not. Two lines of research support the account: That there is an initial bias towards believing early in the process (a temporal effect), and that cognitive load increases the truth bias (an effort effect; Gilbert, 1991; Gilbert et al., 1990, 1993; Mandelbaum, 2014). Our findings challenge this account.

The informed Cartesian account was recently forwarded as an alternative (Street & Richardson, 2015a). The authors claimed that raters made use of context-relevant

knowledge to make the best guess when unsure. They showed that the temporal effect is better explained by an informed Cartesian account. Here, we considered whether a modified Spinozan account is a viable alternative by examining the effort effect in conditions that are favourable to a Spinozan account (Gilbert et al., 1990). We found raters were more likely to say they believed a speaker was telling the truth if there was a cognitive load, but only if they were forced into making a lie-truth response. Those who could explicitly indicate their uncertainty, and no longer needed to guess, did not show this bias whether they were under load or not. The findings support an informed Cartesian account: When unsure (due to load), the truth bias reflects the best guess when required to judge. These findings also dovetail with the predictions of ALIED theory (Street, 2015), which claims that people make context-informed guesses under uncertainty.

In recent years, there have been a number of challenges to the Spinozan position. For instance, a number of studies show that people can be biased to believe or disbelieve early in the judgement and with little effort, depending on prior experience and expectations (Deutsch et al., 2009; Hanks, Mazurek, Kiani, Hopp & Schadlen, 2011) and the nature of the information available (Hasson et al., 2005; Nadarevic & Erdfelder, 2013). And yet other studies find support for the Spinozan position (for an up to date review, see Mandelbaum, 2014). The current study offers a way to reconcile these conflicting streams of research by showing how performance can appear both Spinozan and non-Spinozan. Critically, biased performance depends on whether people are forced to make a judgement when uncertain. That is, it is not the structure of the mind per se, but rather the structure of the world that results in a bias to believe (or disbelieve) (see Street, 2015).

What does this mean for the Spinozan account? The account rests largely upon two sources of evidence. First, people show a bias to believe when they do not have sufficient cognitive resources to move on to the evaluative stage that allows for disbelief. Second, people show a bias to believe when under time pressure, and require more time to create 'false tags' and thence disbelieve. Here, we show that the first prediction can be better explained by an informed Cartesian account: Cognitive load creates uncertainty and results in an attempt to make an informed, albeit biased, guess. The second prediction can

also be better explained by an informed Cartesian account – people rely on context to make an informed judgement early in the decision process (Street & Richardson, 2015a, b).

A devoted advocate of the Spinozan position may want to move towards claiming the account explains only certain limited phenomena.[5] For instance, while it may be that the Spinozan account cannot explain social judgements (and ignoring for the moment the fact that the theory explicitly proposes that the truth bias observed in lie detection studies is the result of the Spinozan structure of the mind; Gilbert, 1991; Gilbert et al., 1990, 1993), it may still explain judgements of written propositions. When proposing the Spinozan account, Gilbert et al. (1990) took the position that the Spinozan mind must influence belief in its widest form – whether a seemingly deliberative evaluation of belief, or an automatic accepting. For example, Gilbert (1991) explains that the visual input from our eyes is a form of belief, insomuch as we 'believe our eyes'. Thus, the Spinozan account claims that visual processing is a form of belief. Gilbert (1991) also gave considerable space to discussing beliefs that may be considered to be more deliberative, including lie detection, the truth bias, and even court sentencing judgements where participants were explicitly expected to deliberate on their sentencing (see Gilbert et al., 1993). In fact, they offered empirical evidence that the truth bias in lie detection tasks is a result of the Spinozan mind (Gilbert et al., 1990, Study 2). Evidently, any claims that the Spinozan mind does not attempt to explain social judgements such as lie-truth judgements are not at all consistent with the empirical evidence or the theory itself (see Gilbert, 1991).

Contracting the account to only explain particular types of belief (e.g., only isolated propositions) seriously undermines the account. After all, the account proposes it is the structure of the mind, not the nature of the task, that results in an automatic bias to believe. If it is just the task set-up that causes the bias, then no strong claims can be made to the cognitive structure. It is important, then, that the account can generalize to a variety of truth-belief situations, and that the task set-up is not the cause of the bias.

Yet the task set-up *does* seem to be the responsible for the truth biases previously observed. This undermines the claim that biases are the result of a Spinozan structure of

[5] An anonymous reviewer of this manuscript took this position.

the mind. For instance, that negated propositions (e.g., 'The star is not above the cross') take longer to process than affirmed propositions (e.g. 'The star is above the cross'; Clark & Chase, 1972; Mayo et al., 2004; Zwaan, Stanfield, & Yaxley, 2002) has been considered 'a theoretical coup for the Spinozan', and that 'the quintessential anti-Cartesian experimental paradigm is one that exploits asymmetries in people's memory of truths and falsehoods' (Mandelbaum, 2014). But research has shown that this temporal effect can be better explained by the greater informativeness of affirmed compared to negated propositions (Anderson, Huette, Matlock, & Spivey, 2009; Glenberg & Robertson, 1999; Hasson et al., 2005) – even this theoretical coup results from the nature of the information available, rather than the architecture of the mind.

The informed Cartesian and the Spinozan accounts differ not only in the process that causes the bias, but also in the conceptualization of the bias itself. The Spinozan account views the bias as a fixed constant built into the cognitive system, and as a result, it can cause *errors* in understanding, particularly if the information should not be believed. In contrast, the informed Cartesian account views the bias as a flexible adaptation to the current context (see Glöckner, Hilbig, & Jekel, 2014; Simon, 1990; Street, 2015), as informed by prior experience, and as *aiding* the judgement rather than hindering it (for an overview of generalized heuristic rules improving judgments, see Gigerenzer et al., 1999; Gilovich, Griffin, & Kahneman, 2002). The informed Cartesian account is thus highly consistent with the ALIED theory (Street, 2015).

It is adaptive to believe people are telling the truth because people usually do tell the truth (DePaulo et al., 1996; Grice, 1975; McNally & Jackson, 2013). If someone had to guess whether a statement was a lie or truth, even before hearing any information, the best strategy would be to guess people are being honest (see Nadarevic & Erdfelder, 2013). But to be truly adaptive, the direction of the bias should reflect the current context, such that if the context suggests people are more likely to lie, raters will show a lie bias. Making raters suspicious or distrustful increases the likelihood of lie judgements (Mayo et al., 2004; Richter et al., 2009; see also Levine et al., 2010; Sperber et al., 2010), and when people expect to hear mostly lies, they show a lie bias (Blair, 2006; Masip, Alonso, Garrido, & Herrero, 2009; Street & Richardson, 2015b; see also Meissner & Kassin, 2002). This is consistent with the predictions of ALIED theory (Street, 2015).

Before concluding, it is important to discuss the extent to which button responses reflect unconscious processes. A Spinozan may argue that even though participants indicate uncertainty at a conscious and explicit level, at the unconscious level they are still biased towards believing. Put another way, our public measure may not be sensitive enough to assess the private cognitive processes at work.

We have two responses to this. Caution must be heeded when dismissing findings based solely on the possibility that what people can be observed to be doing at their fingertips is not necessarily a direct reflection of what processing is occurring in the cognitive black box. Such a strong critique leaves us in the position of having to dismiss the whole body of behavioural psychological research. Of course, the work supporting the Spinozan mind account must also be dismissed, because it too is based on behavioural, explicit judgements (e.g., Gilbert et al., 1990, 1993). In such a position, we have no grounds to support a two-process Spinozan account (an automatic biased process and a deliberative unbiased process) over a single-process informed Cartesian account (a conscious process that incorporates context). Indeed, in the absence of evidence one would imagine that positing an additional process that can neither be controlled by the thinker nor measured by the scientist is less parsimonious than claiming that only a single measurable process exists (see Street & Vadillo, in press).

Even if one were to dismiss behavioural data in this way, the current study still undermines the Spinozan account. Researchers before us have presented a methodology that supposedly evidences unconscious and automatic thinking - namely the use of a cognitive load to prevent conscious processing. It is this same method that we used here. Our study shows that the responses on this task are guesses when uncertain. Put another way, it is not particularly important to us whether the responses we obtained are conscious or unconscious. What is important is that in either case, the Spinozan bias found by Gilbert et al. (1990) is only evident when the task is set up in a particular way – using the forced-choice set-up as devised by Gilbert et al. Evidently, the responses are dependent on the structure of the task, not the structure of the mind. At the very least, our findings suggest that the evidence so far can be explained by a conscious process, and do not require the assumption of an additional hidden and uncontrollable process that leads people astray.

Conclusion

The asymmetry in making more truth than lie judgements has been thought to be evidence that people are Spinozans. The informed Cartesian account explains how the Spinozan mind is only illusory. There have been two major forms of asymmetry evidence supporting the Spinozan position. First, there is a bias towards believing rather than disbelieving early in the process (a temporal effect; Gilbert et al., 1990, 1993). However, this can be better thought of as a best guess when unsure, as proposed by the informed Cartesian account (Street & Richardson, 2015a, b). Second, there is a bias towards believing when participants are put under a cognitive load (an effort effect; Gilbert et al., 1990). The current study used a paradigm favourable to the Spinozan account to explore the effort effect. We replicated the early bias towards believing if there is a cognitive load, but only if participants were forced to make a lie-truth judgement. If they were not forced to guess, but could explicitly say they were unsure, there was no such early bias. The effort effect can be better explained as a best guess when unsure, as the informed Cartesian account proposes. These findings are consistent with the ALIED account of the truth bias (Street, 2015). We conclude that people do not have an automatic bias towards naïvely believing everything they hear. Rather, they make smart informed guesses when they are unsure.

References

Anderson, S. E., Huette, S., Matlock, T., & Spivey, M. J. (2009). On the temporal dynamics of negated perceptual simulations. In F. Paril, V. Tobin, & M. Turner (Eds.), *Meaning, Form & Body* (pp. 1-20). Stanford, CA: CSLI.

Birmingham, E., Bischof, W. F. & Kingstone, A. (2008). Gaze selection in complex social scenes. *Visual Cognition, 15*, 341-355. doi: 10.1080/13506280701434532.

Birmingham, E., Bischof, W. F. & Kingstone, A. (2009). Saliency does not account for fixations to eyes within social scenes. *Vision Research, 49*, 2992-3000. doi: 10.1016/j.visres.2009.09.014.

Blair, J. P. (2006). From the field: Can detection of deception response bias be manipulated? *Journal of Crime and Justice*, *29*(2), 141-152. doi: 19.1080/0735648X.2006.9721652.

Clark, H. H., & Chase, W. G. (1972). Process of comparing sentences against pictures. *Cognitive Psychology*, *3*(3), 472-517. doi: 10.1016/0010-0285(72)90019-9.

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), 979-995. doi: 10.1037/0022-3514.70.5.979.

Deutsch, R., Kords-Freudinger, R., Gawronski, B., & Strack, F. (2009). Fast and fragile: A new look at the automaticity of negation processing. *Experimental Psychology*, *56*(6), 434-446. doi: 10.1027/1618-3169.56.6.434.

Gigerenzer, G., Todd, P. M., & The ABC Research Group (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.

Gilbert, D. T. (1991). How mental systems believe. *American Psychologist*, *46*(2), 107-119. doi: 10.1037/0003-066X.46.2.107.

Gilbert, D. T., Krull, D. S., & Malone, P. S. (1990). Unbelieving the unbelievable: Some problems in the rejection of false information. *Journal of Personality and Social Psychology*, *59*(4), 601-613. doi: 10.1037/0022-3514.59.4.601.

Gilbert, D. T., Tafarodi, R. W., & Malone, P. S. (1993). You can't not believe everything you read. *Journal of Personality and Social Psychology*, *65*(2), 221-233. doi: 10.1037/0022-3514.65.2.221.

Gilovich, T., Griffin, D. W., & Kahneman, D. (2002). *Heuristics and biases: The psychology of intuitive judgment*. Cambridge: Cambridge University Press.

Glenberg, A. M., & Roberston, D. A. (1999). Indexical understanding of instructions. *Discourse Processes*, *28*(1), 1-26. doi: 10.1080/01638539909545067.

Glöckner, A., Hilbig, B. E., & Jekel, M. (2014). What is adaptive about adaptive decision making? A parallel constraint satisfaction account. *Cognition*, *133*, 641-666. doi: 10.1016/j.cognition.2014.08.017.

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.

Hanks, T. D., Mazurek, M. E., Kiani, R., Hopp, E., & Shadlen, M. N. (2011). Elapsed decision time affects the weighting of prior probability in a perceptual decision task. *Journal of Neuroscience*, *31*(7), 6339-6352.

Hasson, U., Simmons, J. P., & Todorov, A. (2005). Believe it or not: On the possibility of suspending belief. *Psychological Science*, *16*(7), 566-571. doi: 10.1111/j.0956-7976.2005.01576.x.

Jeffreys, H. (1961). *Theory of probability*. Oxford: Oxford University Press.

Levine, T. R. (2014). Truth-default theory (TDT): A theory of human deception and deception detection. *Journal of Language and Social Psychology*, *33*(3), 378-392. doi: 10.1177/0261927X14535916.

Levine, T. R., Kim, R. K., & Blair, J. P. (2010). (In)accuracy at detecting true and false confessions and denials: An initial test of a projected motive model of veracity judgments. *Human Communication Research*, *36*, 81-101. doi: 10.1111/j.1468-2958.2009.01369.x.

Lick, D. J., & Johnson, K. L. (2014). "You can't tell just by looking!": Beliefs in the diagnosticity of visual cues explain response in social categorization. *Personality and Social Psychology Bulletin*, *40*, 1494-1506. doi: 10.1177/0146167214549323.

Mandelbaum, E. (2014). Thinking is believing. *Inquiry: An Interdisciplinary Journal of Philosophy*, *57*(1), 55-96. doi: 10.1080/0020174X.2014.858417.

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. doi: 10.1002/acp.1535.

Masip, J., Garrido, E., & Herrero, C. (2006). Observers' decision moment in deception detection experiments: Its impact on judgment, accuracy, and confidence. *International Journal of Psychology*, *41*(4), 304-319. doi: 10.1080/00207590500343612.

Mayo, R., Schul, Y., & Burnstein, E. (2004). "I am not guilty" vs "I am innocent": Successful negation may depend on the schema used for its encoding. *Journal of Experimental Social Psychology*, *40*(4), 433-449. doi: 10.1016/j.jesp.2003.07.008.

McNally, L., & Jackson, A. L. (2013). Cooperation creates selection for tactical deception. *Proceedings of the Royal Society B: Biological Sciences*, *280*, http://dx.doi.org/10.1098/rspb.2013.0699.

Meissner, C. A., & Kassin, S. M. (2002). “He’s guilty!”: Investigator bias in judgments of truth and deception. *Law and Human Behavior*, *26*(5), 469-480. doi: 10.1023/A:1020278620751.

Morey, R. D., & Rouder, J. N. (2015). *BayesFactor* (Version 0.9.11-3) [Computer software]. Retrieved from http://www.jasp-stats.org.

Nadarevic, L., & Erdfelder, E. (2013). Spinoza’s error: Memory for truth and falsity. *Memory and Cognition*, *41*, 176-186. doi: 10.3758/s13421-012-0251-z.

R Development Core Team (2011). *R: A Language and Environment for Statistical Computing* [Computer software]. Retrieved from http://wwww.r-project.org.

Richter, T., Schroeder, S., & Wöhrmann, B. (2009). You don't have to believe everything you read: Background knowledge permits fast and efficient validation of information. *Journal of Personality and Social Psychology*, *96*(3), 538-558. doi: 10.1037/a0014038.

Rouder, J., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian *t* tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*(2), 225-237. doi: 10.3758/PBR.16.2.225.

Schroeder, S., Richter, T., & Hoever, I. (2008). Getting a picture that is both accurate and stable: Situation models and epistemic validation. *Journal of Memory and Language*, *59*, 237-255. doi: 10.1016/j.jml.2008.05.001.

Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology*, *41*, 1-19. doi: 10.1146/annurev.ps.41.020190.000245.

Skurnik, I., Yoon, C., Park, D. C., & Schwarz, N. (2005). How warnings about false claims become recommendations. *Journal of Consumer Research*, *31*(4), 713- 724. doi: 10.1086/426605.

Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D. (2010). Epistemic vigilance. *Mind and Language*, *25*(4), 359-393. doi: 10.1111/j.1468-0017.2010.01394.x.

Street, C. N. H. (2015). ALIED: Humans as adaptive lie detectors. *Journal of Applied Research in Memory and Cognition*, *4*(4), 335-343. doi: 10.1016/j.jarmac.2015.06.002.

Street, C. N. H., & Richardson, D. C. (2015a). Descartes versus Spinoza: Truth, bias, and uncertainty. *Social Cognition*, *33*(3), 227-239. doi: 10.1521/soco.2015.33.2.2.

Street, C. N. H., & Richardson, D. C. (2015b). Lies, damn lies, and expectations: How beliefs about the base rate inform lie-truth judgments. *Applied Cognitive Psychology*, *29*, 149-155. doi: 10.1002/acp.3085.

Street, C. N. H., Tbaily, L., Baron, S., Khalil-Marzouk, P., Hanby, B., Wright, K., & Richardson, D. C. (2011, April). Bloomsbury deception set. *Twentieth Annual Conference of the British Psychological Society Division of Forensic Psychology, Portsmouth, UK*.

Street, C. N. H., & Vadillo, M. A. (in press). Can the unconscious boost lie detection accuracy? *Current Directions in Psychological Science*.

Unkelbach, C. (2007). Reversing the truth effect: Learning the interpretation of processing fluency in judgments of truth. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(1), 219-230. doi: 10.1037/0278-7393.33.1.219.

Zellner, A., & Siow, A. (1980). Posterior odds ratios for selected regression hypotheses. In J. M. Bernardo, M. H. DeGroot, D. V. Lindley, & A. F. M. Smith (Eds.), *Bayesian Statistics: Proceedings of the First International Meeting* (pp. 585-603). Valencia: University of Valencia Press.

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally represent the shapes of objects. *Psychological Science*, *13*(2), 168- 171. doi: 10.1111/1467-9280.00430.

Appendix A: Bayes Factor

Null findings under null hypothesis significance testing do not give positive evidence of a lack of an effect, only that an effect (whether present or absent) was not observed. It was predicted that there should be a lack of bias in the LTU unforced choice condition, placing the hypothesis on the side of the null. To negotiate this issue, a Bayes factor was calculated. The Bayes factor pits two competing models against each other (e.g., a null and alternative hypothesis). As a result, it can either find (a) positive evidence in favor of the null, (b) positive evidence in favor of the alternative, or (c) the data are inconclusive, e.g. due to a lack of power. In this way we can ask whether our data show evidence of a lack of bias, rather than merely an inconclusive result that may be due to a lack of power.[6]

The BayesFactor package version 0.9.4 (Morey & Rouder, 2013) designed for R (R Development Core, 2011) was used to calculate the Bayes factor. A prior Cauchy distribution with a scaling factor of $r = \sqrt{2}/2$ was used over the standardized effect sizes, as recommended by Rouder, Speckman, Sun, Morey and Iverson (2009; see also Jeffreys, 1961; Zellner & Siow, 1980).[7]

Results

A Bayes factor was calculated to further examine the lack of bias difference in the LTU condition when judging items learnt under cognitive load versus items learnt without load. The odds of no difference in bias to there being a difference in bias are 10.88 to 1. Put another way, in order to prefer the Spinozan interpretation of there being a bias difference, we would need prior odds favoring it of greater than 10.88. Thus the data show strong support for there being no difference in bias between interrupted and uninterrupted items in the LTU condition.

---

[6] As a (very) rough guide, Bayes factors between 1 and 3 are generally considered to be indicative of a lack of power. However, we caution against using these arbitrary cut-off values as a categorical boundary in the way that NHST has adopted p=0.05 as its categorical boundary of significance.

[7] This prior also has the advantages of being readily computable and gives a stable integration of the likelihood.

Appendix B: Three-Way ANOVA on PTJ

A 2 (response condition: LT or LTU, between subjects) x 2 (speaker veracity: lie or truth, within subjects) x 2 (cognitive load: load or no load during learning, within subjects) mixed ANOVA was conducted on the PTJ. The critical analyses related to the Spinozan and Cartesian accounts are presented in the main body of the text. Here, unplanned comparisons for the remaining main effects and interactions are reported. We recommend caution in interpreting these unplanned analyses.

The main effect of interruption was negligible, F (1, 79) = 1.43, p = .236, η_p^2 = 0.02. The interaction between response condition (LT or LTU) and speaker veracity was also of a very small effect size, F (1, 79) = 1.42, p = .236, η_p^2 = 0.02, as was the interaction between cognitive load and speaker veracity, F (1, 79) = 0.73, p = .396, η_p^2 = 0.01. Finally, the three-way interaction also had a negligible effect size, F (1, 79) = 0.07, p = .787, $\eta_p^2 < 0.01$.