How contemporary theory informs lie detection accuracy and bias

Greg Best, Jonathan Hodgeon, Chris N. H. Street

aDepartment of Psychology, University of Huddersfield, Queensgate, Huddersfield, HD1 3DH, University Kingdom
bEconomic Crime Unit, North Yorkshire Police, Alverton Court, Crosy Road, Northallerton, DL6 1BF, United Kingdom

KEYWORDS

Accuracy
Adaptive lie detector
Bias
Deception detection
Truth default theory

ABSTRACT

Lie detection research has reached a stage where theory building is possible. We believe practitioners should contribute to theory as it is being developed. With this in mind, we briefly review contemporary theories - namely the Adaptive Lie Detector theory and Truth Default Theory - and consider the practical implications for reducing bias and increasing accuracy when making lie-truth judgements. There are practical issues that theory has yet to address, which are briefly considered. This article intends to spotlight the on-going academic work in lie detection in hope that practitioners will feedback to researchers about how theory may relate to their work, and in turn help steer the direction of future research.

Attempts to detect lies can be traced back at least as far as ancient Greece (Trovillo, 1939), but to date, no reliable method has been discovered. When novel methods are explored, such as the use of fMRI to measure brain activity, countermeasures are developed that make them useless (see Ganis, Rosenfeld, Meixner, Kievit & Schendan, 2011). Countermeasures like these have the effect of nullifying the original discovery and taking us back to the drawing board to invent a new means of lie detection. When a countermeasure is developed, we should not have to wait for the next spark of creative inspiration. A theoretical framework can readily offer identifiable paths to developing a new intervention or bolstering the original intervention to nullify the countermeasure. In this article, we consider how two contemporary theories of lie detection speak to practical matters of detecting lies.

Theorising in the area is still nascent, and so while we consider the implications of theory for practitioners, we would encourage caution in adopting any suggestions here until further research has more rigorously tested the theories. This forward-looking article aims to spotlight the academic work on-going in the area with the hope that practitioners will feed back to researchers about how theoretical work may relate to their practice, and to work with researchers to help steer the direction of future research.

The Adaptive Lie Detector and Truth Default Theories

In this section we briefly outline the Adaptive Lie Detector theory (ALIED: Street, 2015) and Truth Default Theory (TDT: Levine, 2014). Later we will consider what both can tell us about accuracy and bias in lie detection.

ALIED (Street, 2015) proposes that people make informed judgements of whether someone is lying or telling the truth. If there are reliable cues to deception or honesty that directly relate to the claim being made, raters can use those cues to achieve high accuracy. Information that directly relates to a statement might include body language or the words used, but these are typically not reliable (DePaulo et al., 2003). ALIED calls these ‘individuating cues’ because they individuate that statement from other statements. More reliable individuating cues might include physical evidence such as a time-stamped photograph of being in the
location that was claimed in the statement.

In the absence of reliable individuating cues that directly relate to the claim, people may instead make an inference based on their beliefs about the situation more generally, which ALIED calls ‘context-general information’. This information does not directly relate to the claim being made, but instead generalises across statements. For example, in our daily lives most people tell the truth most of the time (DePaulo et al., 2003), which leads to greater reliance on context-general information such as ‘most people tell the truth’. If people had a different generalised belief about a situation (e.g., that prisoners tend to lie), then ALIED would predict a bias to disbelieve others, called a lie bias. Such findings have been observed (e.g., Bond, Malloy, Arias, Nunn & Thompson, 2005).

In contrast, TDT (Levine 2014) suggests that people do not vicariously adapt to the situation, but rather operate on the default presumption that what another person says is truthful. A ‘trigger’ is needed to be released from this bias. These triggers may include (i) a belief that the speaker has a motive to lie, (ii) behaviour and an appearance that gives the impression of dishonesty, (iii) inconsistencies detected within a statement (e.g., contradicting what was said earlier), and (iv) inconsistencies detected between a statement and known facts (e.g., contradicting a photograph of them at the scene: Levine, 2014). If a lie detector no longer relies on their truth-default, TDT suggests that people will engage in more evaluative reasoning. A lie judgement will be reached only if there is sufficient evidence to justify it, otherwise people fall back to assuming the speaker is telling the truth. Sufficient evidence matches those triggers that lead to an abandonment of the truth-default, listed above.

ALIED theory reviews evidence that challenges a defaulting account (Street, 2015). We are not aware of a positional paper that challenges the tenets of ALIED theory. While the two accounts take different stances (flexible and adaptive reasoning versus relying on default beliefs), the practical implications resulting from each have substantial overlap. This article considers the implications of these theories for practical lie detection and where they are currently lacking.

Biases in Lie Detection

Biases in lie detection have previously been considered an error in judgement formation (Buller & Burgoon, 1996; Gilbert, 1991; O’Sullivan, 2003). Yet ALIED and TDT agree that the truth bias (i.e., judging claims to be truthful more often than deceptive) reflects the true state of the world: That most people are honest most of the time (DePaulo et al., 1996; Halevy et al., 2014; Serota et al., 2010). Although lay people are ordinarily biased to believe claims, when made to feel suspicious (Blair, 2006; DePaulo & DePaulo, 1989; Masip, Alonso, Garrido & Herrero, 2009) they can be biased to disbelieve. Thus, according to ALIED, the tendency to believe or disbelieve depends on one’s generalised beliefs about that context (in the absence of reliable cues).

While ALIED argues that the bias to believe is only present in those situations where people tend to tell the truth, TDT argues that it is an ever-present “passive presumption” or a “fall back cognitive state after a failure to obtain sufficient… evidence” (Levine, 2014, p.380). It is not clear how TDT explains findings that people can be biased to disbelieve in some situations (e.g., when interacting with car salespeople, prisoners, or criminal suspects: e.g., DePaulo & DePaulo, 1989; Masip et al., 2009). Presumably, it would require that people have a low threshold for giving up their truth-default and then engage in an evaluation that requires only little evidence that a person is lying. However, it is worth noting that elsewhere the author has argued that there is not sufficient evidence that people can be biased to disbelieve (Levine et al., 2014).

ALIED argues that bias is typically a ‘good guess’ when there are no good cues to deception, and that people rely on their knowledge of the situation more generally to fill in for the absence of reliable information. If ALIED is correct, reducing the bias may mean encouraging people to recognise that reliable individuating cues are absent and to not make a judgement in that situation. This has seen some support (Street & Kingstone, 2016; Street & Richardson, 2015; see also Nadarevic & Erdfelder, 2013).

Rather than removing the bias to believe, practitioners may wish to shift towards a bias to disbelieve. In reviews of child deaths and serious injuries in England and Wales, 75% involved the caregiver deceiving social workers into believing that they were not neglecting their child (Brandon et al., 2009). There have been calls within the social work community to place less belief in parents’ claims and to ‘think the unthinkable’ or maintain a ‘respectful uncertainty’ by considering the
possibility of deception (Department of Health, 2003; Naqvi, 2013). In these situations, it may be prudent to have a bias to disbelieve, or a ‘lie bias’ (Meissner & Kassin, 2002). ALIED claims that a shift towards a lie bias may be seen when people hold the belief that most people lie in this context (but only if there are no reliable individuating cues available to otherwise guide the judgement). Thus, a lie bias may result from providing training to focus on the possibility of deception, supported in research (Blair, 2006; Masip et al., 2009; Toris & DePaulo, 1985; see also Masip & Herrero, 2017). Importantly, according to ALIED this should be context-dependent, so that training to detect lies in a police setting should lead an officer to be less likely to believe statements in work life but not in their home life (see Masip & Herrero, 2017).

TDT claims that people will abandon a truth belief if there is a trigger, such as the speaker producing behaviours that look deceptive (even if they are not related to deception). This results in a more evaluative decision. There is an undefined threshold of what makes something sufficient evidence to be considered deceptive, but a lie bias would be observed if people were more readily willing to interpret information as evidence of deception.

Lie Detection Accuracy: Why Is It Poor and What Can We Do?

People are generally poor at detecting deception and obtain accuracy rates around the level of chance (Bond & DePaulo, 2006). This is likely due to the fact that liars do not typically display reliable cues to deception (Burgoo, 2018; DePaulo et al., 2003; Van Der Zee, Poppe, Taylor & Anderson, 2015), if any at all (Levine, 2010). While ALIED has implications for reducing bias, TDT has more explicit recommendations for improving accuracy.

Both ALIED and TDT argue that people should rely less on nonverbal behaviours, which are unreliable (DePaulo et al., 2003), and instead focus on using more reliable information, such as comparing what is said to known facts (e.g., physical evidence: Blair, Reimer & Levine, 2018; Granhag, Strömwall, Willén & Hartwig, 2012; Nahari, Vrij & Fisher, 2014). The danger, according to ALIED, is that a motivated liar can exploit this (Street, 2015). If the liar will have a prolonged interaction with another person (e.g., a player in poker game, a juror at a court case), the liar can produce a behaviour that is present only when they tell the truth. This can later be verified by the person trying to detect the lie by matching the claim against known evidence. In the case of poker, that may be looking at the person’s hand once the hand has been played out. In a court case, that may be checking the claim against physical evidence presented in the case, such as CCTV footage. The lie detector may learn to associate this behaviour with that person’s honesty. At an opportune moment (e.g., when the poker stakes are high, or where physical evidence is likely to be absent, such as the question of whether a rape allegation was consensual or not), the liar may produce the behaviour that has so far been associated with honesty. In situations such as this, decision makers may come to rely on the so-far reliable behaviour that the liar intentionally produced while being honest.

It is in part for this reason that ALIED does not offer an explanation of how to improve accuracy: How accurate people are will depend upon the information they detect and their beliefs about how reliable that information is. Both of these are open to manipulation by the deceiver. However, in general, ALIED argues that the elicitation of reliable clues (e.g., using the verifiability approach, Nahari et al., 2014, or the strategic use of evidence, Granhag et al., 2012) will lead to higher accuracy. This is because ALIED assumes that people are able to detect those cues and will use them instead of less reliable clues (for evidence of this, see for example Bond, Howard, Hutchison & Masip, 2013; Hartwig & Bond, 2011; Levine et al., 2014; Street, Bischof, Vadillo & Ringstone, 2016).

TDT puts forward the ‘content in context’ approach (Blair, Levine & Shaw, 2010) as a means of prompting such information. This approach encourages lie detectors to ignore the visible appearance of dishonesty and instead consider how plausible a claim is, both in terms of the likelihood of the claimed event happening in that situation and how likely the speaker is to have behaved in that way. Implausible details should be judged as lies. Although TDT does not offer instruction on how to incorporate this into a questioning script, it does argue that people can be trained in this approach, and that it results in higher rates of accuracy when compared to passive lie detection (Blair et al., 2010; Levine, Clare, Green, Serota & Park, 2014; although see Levine & McCormack, 2001, for evidence that active questioning can increase the bias to believe rather than increase accuracy).

In brief, both ALIED and TDT are consistent with current approaches that seek to elicit clues from deceivers that allow for a check against reality. However, we note that truth-tellers can also contradict known facts due to memory decay, for instance. The field is as yet unable to avoid the potential for misclassifying liars as truth-tellers and vice versa.
Individual Differences

We are unlikely to find a ‘super-detector’ who can identify deception better than most people (Bond, 2008; Bond & Uysal, 2007). This may not be surprising when one remembers that cues to deception are rare and unreliable – if the information is not available, no degree of perceptual skill will find them. However, individual ability to detect deception can be conceptualised in a variety of ways other than simply an ability to pick up on and effectively use the already available information to make highly accurate judgements.

One issue decision makers face is that they may not know what information is relevant for making a decision. ALIED notes that this is equivalent to the information not being present. A precursor to making an accurate judgement, then, is to know what information is relevant. This is not just a case of understanding that eye contact is not a good cue to deception (DePaulo et al., 2003), for instance. The last two authors are exploring how people interact with SMS text messaging scams. Sophisticated attacks can make their SMS message appear in the same conversational thread as the company they are attempting to simulate – for example, a bank may send you updates about your account, but a hacker can make their message appear in the same SMS thread. The potential to make an accurate judgement is lacking if one is not aware that this is technologically possible.

TDT frames this as a lack of a trigger that leads to the suspicion of deceit, and so assuming no other triggers (indicators of deceit) are present, a mobile phone user would default to trusting the message. ALIED suggests people will look for individuating cues that suggest honesty or deceit. In their absence, context-general beliefs about SMS messages in general will be considered. These beliefs may differ across people. For instance, one might imagine that someone whose work involves preventing fraud might hold a generalised belief that messages in general cannot necessarily be trusted, while a lay user may have a generalised belief that messages in general tend to be genuine.

Some Open Questions

Theorising is still relatively new to the field of lie detection. Both TDT and ALIED are fairly broad and have substantial ambiguity in how the claimed processes are implemented. For instance, ALIED theory allows for almost any generalised belief to be considered as context-general information (although see Peebles & Street, 2018, for a computational model implementing ALIED that requires precise definition), while TDT notes that people abandon a truth-default belief when there is information that leads to the suspicion of deception, which flirts with being tautological. The aim of this section is neither to dismiss these theories nor undermine their authors. To progress in our understanding, we must attempt to falsify what we believe we know. This section aims to briefly cover areas that ALIED and TDT do not consider, both from a theoretical and practical viewpoint. We hope that this encourages active challenging of both theoretical perspectives and in turn creates a deeper understanding of lie detection.

Aside from the lack of precision, another question that we feel is still somewhat unaddressed is whether people weigh up the costs of being deceived against the effort that would need to be invested to verify a lie. Imagine a friend telling you that they saw the remains of a WWII fighter plane when they were on holiday in France. The cost of not detecting this possible lie may not justify the effort needed to find reliable individuating cues (e.g., verifying that the plane in question does indeed exist). TDT takes the position that people assume the truth until some information triggers an evaluation to consider the possibility of deception. Similarly, ALIED implicitly assumes people rely on context-general information ordinarily unless a reliable individuating cue makes itself available (e.g., a photograph of the person in front of the plane as a reliable cue to honesty, recalling from memory contradictory stories from other people). In day-to-day contexts, there are good reasons to believe people tend to tell the truth - e.g., that most people actually do tell the truth more often than they lie (DePaulo et al., 1996). Where ALIED differs from TDT is when the context changes. In other situations, such as in a prison, context-general information may be that most prisoners tend to lie. Rather than always assuming honesty, ALIED predicts that prisoners will be biased towards believing others will be dishonest (Bond et al., 2005). Whatever the case, neither theory tackles the issue of whether people trade off detection effort for the potential cost of being deceived.

Conclusion

Contemporary theory has taken two different perspectives. ALIED claims that people fluidly adapt to the context and attempt to use the more reliable information available. TDT meanwhile claims that people default to believing others and will only consider the possibility of deception if there are triggers that raise suspicion (regardless of
whether the trigger is a reliable indicator of deception or not). Yet in practice, both accounts offer similar practical solutions. To improve lie detection accuracy, cues to deception need to be elicited, such as checking claims against physical evidence. To reduce bias, ALIED suggests that people need to recognize the lack of reliable information and make an effort to not rely on context-based information, although relatively little testing has been carried out. TDT does not offer suggestions for reducing bias, as far as we can tell. We have also highlighted that both accounts are rather broad and need refining, and that there are concepts that are not addressed by the theories at all, such as whether people weigh up the cost of being deceived against the effort of looking for reliable information. These gaps, we hope, show that there is scope for practitioners to inform researchers about key questions and issues from their line of work that need investigating. We would welcome practitioners reaching out to us and engaging in these discussions.

References


---

1 We would like to thank The AntiSocial Engineer for this example.