SUMMER 2022 / ISSUE 14

CREST SECURITY REVIEW

ISSN 2398-0540 (PRINT)
ISSN 2398-0559 (ONLINE)

Trust

CONTENTS

Highlights

RECOVERING FROM FAILURE

What can organisations do to repair trust? – p20

MY LIFE IN YOUR HANDS

Is it a problem to blindly trust AI? What might happen if you do? – p22

TRUST

CREST SECURITY REVIEW
Editor – Rebecca Stevens
Guest Editor – Prof. Stacey Conchie
Editorial assistant – Sally Bolton
Illustrator & designer – Rebecca Stevens
To contact *CREST Security Review* email
csr@crestresearch.ac.uk

PAST ISSUES

To download (or read online) this issue, as well as past issues of *CREST Security Review*, scan the QR code or visit our website:
crestresearch.ac.uk/magazine

FROM THE EDITOR

Trust permeates most aspects of our lives and allows us to function on a daily basis. We trust the food we order has not been tampered with. We trust a vet to care for our sick pet. We trust our house to keep us warm, safe, and dry. Security contexts are no exception. Trust offers a mental shortcut that allows us to make immediate (and at times automatic) decisions. In some cases, this is useful as it frees up thinking space to focus on other tasks. In other cases, trust-as-a-heuristic, can be problematic.

This issue of *CREST Security Review* brings together articles that consider how trust can help security, but also how it can create security risks when things go awry or when we rely on it too much.

Guest editor, Professor Stacey Conchie, provides an overview of the articles on page 4 and couches this in a summary of what we mean by trust and what defines its existence.

As in every issue, we highlight a couple of pieces of research away from our main focus topic. Dr Olivia Brown reports on her research looking at the relationship between digital data and the risk of offline action in right-wing terrorism (page 28). Dr Ben Lee introduces us to Siege Culture and explains how this has underpinned many of the recent counter terrorism cases linked to the extreme-right in the UK (page 30).

You can find the research that underpins all our articles in the 'Read More' section on page 32. We have also listed (page 35) some of our previous CSR articles that touch on trust (this list is not exhaustive). As always, we value your feedback and welcome your suggestions for topics or research that you would like to see featured in future *CSR*s. You can send your comments to me at b.stevens@lancaster.ac.uk or can use our anonymous online questionnaire that can be accessed through the QR code (right).

Rebecca Stevens
Editor, *CSR*

CREST SECURITY REVIEW FEEDBACK

CSR's goal is to present informative, world-leading research on security threats in an accessible and engaging format. We hope you enjoy reading *CSR*, and would value your feedback on how we're doing.

To make feedback quick (and anonymous), we have set up an online questionnaire. This can be accessed via the QR code or via this link:

https://bit.ly/3PbwCE6

This questionnaire lists all issues of *CSR* with 3 questions next to each. Please only respond to those issues you have read.

Once again, we wouild be extremely grateful for your honest feedback and thoughtful suggestions, which will help us to continue to improve *CSR*.

Thank you.

STACEY CONCHIE

TRUST IN SECURITY CONTEXTS

Guest editor Professor Stacey Conchie provides an overview of the articles focusing on our special topic of trust.

"To trust, or not to trust, that is the question."

When we interact with a person, group, organisation or system, we may ask this question. We may not consciously verbalise this question and nor will we stick with a relationship because of a fear of what comes if we leave through distrust. Yet, the opening line of Hamlet's soliloquy captures the process that people go through when deciding whether to join, remain or exit a relationship.

What does it mean to trust? When a person trusts another (the trustee), they will accept vulnerability by relying on the trustee to do something of value, which affects them, yet which they have no control over. If we can predict the trustee's actions, then the situation does not call for trust. A person often has more to lose from trusting and being betrayed than from the gains of trust being fulfilled. A covert source may gain financially if their handler is trustworthy. However, if their handler is untrustworthy and betrays them, the consequence may be imprisonment or a threat to their life. For this reason, trust is a risky business.

A person often has more to lose from trusting and being betrayed than from the gains of trust being fulfilled.

Trust (i.e., a willingness to accept vulnerability) is strongly related to a person's beliefs about another's trustworthiness. Indeed, trust and trustworthiness are often synonymous in the literature. Many individual qualities have been proposed to indicate how trustworthy a trustee is. The well-used framework of Mayer and colleagues groups these qualities into those that reflect a trustee's ability, their integrity, and their benevolence. Some researchers go one step further and propose a dichotomy, where ability sits on one side and integrity and benevolence on the other. Beliefs about the former are more rational and objective, the latter more emotional and subjective. As we see in this issue of CSR, both are implicated in security contexts. However, the scale shows a bias towards the subjective end (weighted by integrity) when it comes to shaping behaviour.

Several factors influence trust, including personality, cognitive biases (e.g., stereotypes), similar—past—relationships, gossip, appearance, and direct experience with the trustee. Not all factors are equal in their influence. Nor do they have a prevailing effect. For example, personality (or a person's readiness to trust) is most influential when we meet a potential trustee for the first time, but weakens as we interact with them and observe how they treat others. Trust is not static. The base on which it develops changes and with this, so does its relative influence on risk taking.

This issue of CSR looks at trust in different contexts. The first set of papers look at the role of trust in elicitation. Lina Hillner theorises on the difference between rapport and trust. Anna Leslie and Simon Wells draw on the Eliciting Information Framework to illustrate how this distinction plays out at a practical level. Andreea-Antonia Raducu compares trust against similarity and empathy in the context of a source handler-informant life-cycle.

Stacey Conchie and Paul Taylor show us that not all trust judgements occur at a conscious level. They document studies that show how we might capture people's automatic trust judgements through nonverbal and verbal behaviours.

> **If we can predict the trustee's actions, then the situation does not call for trust.**

We then consider what happens when trust is threatened. Emma Barrett summarises betrayal research and shows how security contexts are hotbeds for their occurrence. One outcome of betrayal is distrust, which in some contexts (e.g., disengagement and deradicalisation) can have positive outcomes (see Morrison et al.), but in others can cause retaliation behaviours that pose security risks, as seen with insider attacks. Rosalind Searle draws on her CREST research to illustrate how trust can be damaged within organisations through poor leadership and how this may be avoided. Steven Lockey discusses how trust may be repaired following a breach (e.g., betrayal), and, similar to the work conducted by Mariam Oostinga on communication errors, shows the important role of an apology. He extends this to illustrate the need for concomitant structural changes when a violation occurs at an organisational level.

Trust not only occurs between people. Ella Glikson illustrates this by summarising research on the role of emotional and rational trust in AI. Paul Taylor discusses the importance of trusting research centres. Finally, Calvin Burns points to the role of trust between organisations (and the many forms trust can take) in a concluding A-Z of trust.

The collection of trust articles in the current (and previous issues of) CSR provide a glimpse into the multi-faceted nature of trust. The coverage is not exhaustive, but all agree on its importance. There is certainly much more to be known about trust in this area and we will continue to see new and innovative work around trust in security contexts as we move forward.

Stacey Conchie is a professor in psychology at Lancaster University and Director of CREST.

Image credit: *ASDF* | stock.adobe.com

LINA HILLNER

RAPPORT AND TRUST: WHAT'S THE DIFFERENCE?

Rapport and trust are not the same: how research is attempting to disentangle these concepts.

WHY IS RAPPORT IMPORTANT?

Researchers and practitioners agree about the importance of rapport for effective information gathering in investigative and intelligence contexts. Rapport concerns the quality of the interviewer-interviewee interaction, which can be characterised in terms of mutual attention, positivity, and connected flow between parties. Rapport-building lies at the heart of non-coercive interviewing approaches and is associated with greater satisfaction with the interview procedure and the interviewer's behaviour, increased information disclosure, and more accurate memory retrieval. The benefits of building rapport have been replicated in laboratory and field research with a range of interviewees, including child and adult witnesses, suspects, cooperative sources, and convicted terrorists. As such, evidence-based interview models recommend rapport-building at the early stages of an interview and highlight the importance of maintaining rapport throughout the interview. Importantly, rapport-building should be viewed as a genuine attempt to connect with the interviewee rather than a transactional interviewing strategy. Insincere attempts to build rapport might backfire and render the interviewee less engaged and less cooperative.

WHAT ABOUT TRUST?

Research on information gathering has given little consideration to the role of trust. Trust is the intention to assume vulnerability based upon the expectation of a positive outcome, and has been shown to reduce conflict and increase cooperation in domains such as teams and negotiations. Establishing trust in the interview room might well yield similar benefits.

The apparent neglect of trust in the information gathering literature might partly be due to the conflation of the concepts of rapport and trust. Scholars and practitioners often use these terms interchangeably, which seems reasonable considering that the few studies investigating the role of trust in an interviewing context yield similar results to those investigating rapport. However, other work suggests that trust is qualitatively different from rapport. Indeed, research examining rapport in professional service contexts has failed to show a statistical relationship with trust. This lack of relationship suggests that rapport and trust may be distinct constructs. Clearly, it is important to start disentangling the individual and conjoint effects of trust and rapport on elicitation outcomes.

RAPPORT AND TRUST: RELATED BUT DISTINCT CONCEPTS

At a theoretical level, both concepts reflect qualities that demonstrate their distinctiveness. Rapport is related to the atmosphere and dynamics of an ongoing interactive event; it characterises the degree to which parties pay attention to one another and the natural flow from one topic to another during their conversation. In contrast, trust reduces the perceived risk associated with the outcome of an interaction. It decreases uncertainty about the other parties' behaviour and is thus concerned with the aftermath of an interaction. Therefore, rapport and trust are related but independent concepts. Their relative importance in security contexts most likely varies from context to context and depends on the nature and length of a relationship. Rapport might be particularly important at the early stages of a relationship when little other information is available. Trust, given its dependency on repeated positive exchanges to develop, might play a stronger role over the longer term.

To illustrate the relative importance of trust and rapport, imagine the following scenario: The intelligence services have approached you (i.e., the source) because you have information that is of interest to them, and you are about to have your first conversation with your contact person (i.e., source handler). You are free to give away or withhold information. The decision to do either will most likely depend upon the quality of the interaction. In the absence of previous experience and trust,

> **The conflation of rapport and trust is an oversimplification that will not prove useful in the long run.**

rapport might be the determinant factor in whether you choose to provide information or not. If you fast forward two years, the relative importance of rapport versus trust might shift. By then, you will have had numerous conversations with your source handler and have grown to trust them. Given the stage of your relationship, a short and awkward conversation might not stand in the way of you providing information. Put differently, the fact that you trust your source handler might compensate for the fact that they are having an 'off' day.

This example raises multiple questions that should be addressed by future research:

1. How does the relative importance of rapport and trust develop over the course of a source-source handler relationship?
2. Does the presence of trust in the approaching institution (i.e., law enforcement or security agency) render rapport superfluous in the early stages of a source-source handler relationship?
3. If trust in the source handler has been violated, will rapport-building attempts still be effective?

A WAY FORWARD

It is increasingly clear that the conflation of rapport and trust in investigative contexts is an oversimplification. Although seemingly useful, this conflation may cause more harm than good as the beneficial effects of rapport might be overestimated while the benefits of trust might never be appreciated. At a minimum, researchers who investigate rapport should consider including measures of trust in their studies and vice versa. This will enable us to examine their relative effects. We also need to test rapport-building and trust-building under different circumstances and at different stages of relationship-building to enable new insights into effective context-specific elicitation tactics. A clear picture of the individual and conjoint effects of rapport and trust on information elicitation can unlock multiple new layers of influence, and result in well-informed advice for practitioners.

Lina Hillner is a PhD student at the University of Portsmouth. Her CREST-funded PhD project focuses on the role of rapport and trust in eliciting information in online contexts.

Image credit: *Andrey Popov* | stock.adobe.com

ANNA LESLIE & SIMON WELLS

EVALUATING TRUST & RAPPORT: A PRACTITIONER'S GUIDE

In Lina Hillner's article (pages 6-7) she lays out the differences between rapport and trust and argues the case for further research to disentangle the two. This article discusses the concepts from a practitioner's perspective and demonstrates how CREST's Eliciting Information Framework can help.

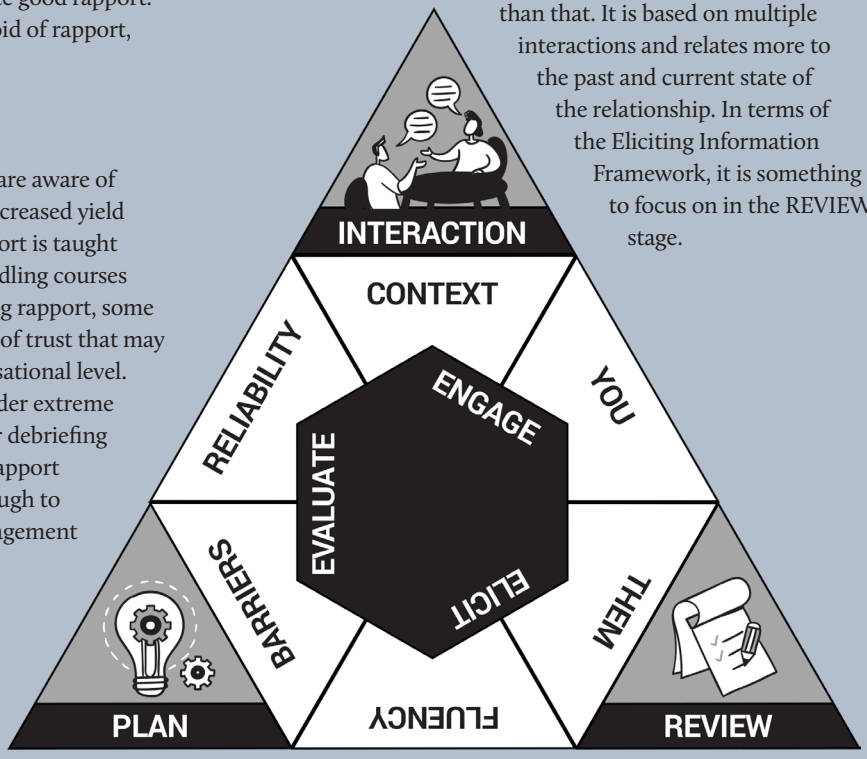
As Hillner says, the concepts of trust and rapport have become conflated, not only by researchers, but also by practitioners. Trust and rapport are separate but related concepts, and it is possible to have one without the other. For example, we all have interactions with people we do not entirely trust (perhaps with certain colleagues), yet our interactions may demonstrate good rapport. Equally, we can have a dreadful interaction, devoid of rapport, with someone we trust deeply.

SO, WHAT?

So, what does this mean for practitioners? Most are aware of the well-established link between rapport and increased yield of credible information. The importance of rapport is taught widely on negotiation, interview and source handling courses (Alison & Alison, 2020). In addition to monitoring rapport, some practitioners focus on trust, including the layers of trust that may (or may not) exist at an individual and an organisational level. However, we recognise that practitioners are under extreme cognitive load when interviewing, negotiating or debriefing (Hanway, 2019). Focussing on whether trust or rapport is present during an interaction may well be enough to reduce listening and thus negatively impact engagement and effective elicitation.

To help, we propose using CREST's Eliciting Information Framework. Both rapport and trust sit under the function of ENGAGE; they are both concerned with having a consistent and positive interaction with the other person as a means to elicit maximum information. Conceptually though, they perhaps relate to different phases of the interaction.

Rapport is all about the INTERACTION; the flow of the conversation, the attention that two parties give each other and whether there exists a genuine desire to connect. Trust is larger than that. It is based on multiple interactions and relates more to the past and current state of the relationship. In terms of the Eliciting Information Framework, it is something to focus on in the REVIEW stage.



TOOLS

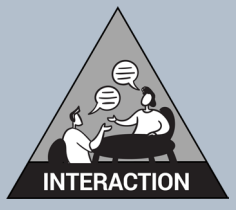
We recommend using the following tools and techniques to help you build rapport and trust.

Plan

We suggest that you plan to build both cognitive and affective trust (Lewis & Wiegert, 1985). **Cognitive trust** in this context is someone's measure of your competence. Many factors feed into this including your appearance, the layout of the room and the credibility of the logistical planning. In order to develop **affective trust**, consider how you will communicate empathy, how you will show that you are interested in them, and how you will demonstrate that you trust them.

To achieve maximum yield of credible information we advocate the use of rapport-building non-coercive techniques. This requires planning; use language in line with your objective, consider what you already know about the individual and how to use this to facilitate their comfort, and rehearse your approach.

Interaction



Engage:

In order to build engagement, be guided by simple acronyms such as OARS:

OPEN-ENDED QUESTIONS

AFFIRMATIONS

RESPONSIVE LISTENING

SUMMARIES

Remember that being in charge and setting an agenda, being frank and forthright, while at the same time social and warm, will also be seen as non-judgemental and is likely to build rapport and increase yield (Alison, Humann & Waring, 2016).

Evaluate:

As already discussed, if you are interviewing on your own then attempts to measure trust and rapport during the interaction may lead to cognitive overload. As an alternative, we suggest that you simply consider whether you are in or out of 'sync'. Signs of being out-of-sync include:

Remember that being in charge and setting an agenda, being frank and forthright, while at the same time social and warm, will also be seen as non-judgemental and is likely to build rapport and increase yield.

- The interviewer/negotiator working harder than the subject;
- Overtalking;
- Too many questions;
- General signs of agitation or anger;
- The subject not engaging and being avoidant.

When you are in-sync, it feels and sounds good, and you are eliciting information that helps your objectives. Put simply, focus on the other party, respond appropriately and encourage someone to say more.

Review

This phase is your real opportunity to review trust and rapport. If you were in-sync and had an increasing level of yield, then you were cooperating and will have had rapport. If that didn't happen, we recommend that you examine rapport (within the interaction), separately from trust (within the relationship), and focus on the problem spaces FLUENCY, BARRIERS, and THEM.

IN CONCLUSION

This article is designed to reduce the confusion between trust and rapport by demonstrating use of CREST's Eliciting Information Framework. Plan for both, and review whether your interaction included rapport and whether your relationship has components of trust. But when engaging, aim to just stay in sync. If you do something which feels uncomfortable, take a breath, assess, and don't be afraid to ask what has changed during the dynamic.

Anna Leslie and Simon Wells are Research to Practice Fellows for CREST. They apply behavioural and social science research to a range of law enforcement, security, and defence issues via training and consultancy.

ANDREEA-ANTONIA RADUCU

TRUST THY ENEMY:
TRUST AND RELATIONSHIP-BUILDING BETWEEN SOURCE HANDLERS AND INFORMANTS

To elicit intelligence from informants, source handlers have to first gain their trust. This takes time and is likely to change as the relationship moves through different stages.

In 2018, informants helped safeguard over 200 people by disrupting terrorism and organised crime (Home Office, 2021). Studies show that effectively eliciting information from informants relies on the development of rapport, trust, effective questioning, and deception detection. Often these studies focus on single-episode interactions (initial encounters between two parties). Less common is for research to consider the ebb and flow of elicitation processes across the life-cycle of the relationship. The importance of considering these changes can be illustrated with trust. Trust sits at the heart of a source-handler-informant relationship as each accepts a level of vulnerability by sharing sensitive information and relying on each other to make decisions that may critically affect them. Having neither oversight nor control of each other's behaviour makes trust crucial. We know that trust is neither uni-dimensional nor static, that it can take different forms and change over-time. An awareness of this is especially important to source handler -informant relationships as these are conceived at the outset as being longer-term partnerships (compared to, say, a police interviewer and witness).

THE LIFE-CYCLE OF A SOURCE HANDLER-INFORMANT RELATIONSHIP: A STAGE MODEL

One way to understand the life-cycle of a source handler-informant relationship is through Knapp's (1978) Staircase Model. This model maps the development, maintenance and dissolution of relationships over ten stages. The first five stages map the coming together of the relationship, and the final five stages map the termination of the relationship.

Let's consider the first 5 stages to understand how trust may play out. Here the relationship progresses through:

1. The Initiating stage: where the source handler and informant have just met and spend time scanning each other, proceeding with caution when interacting;
2. The Experimenting stage: where the pair try to gather information about the other, searching for commonalities and engaging in small talk;
3. The Intensifying stage: where information disclosure has more depth and the source handler and informant develop shared meanings and where possible, experiences;
4. The Integrating stage: where the pair become more synchronised in behaviours and speech patterns, and become increasingly similar (perceptually or actually); and
5. The Bonding stage: an extension of a previous stage, representing a legal commitment in the relationship.

The model predicts that with each stage comes a greater level of self-disclosure. Simply put, a better relationship will result in more actionable information.

WHAT DETERMINES MOVEMENT BETWEEN STAGES?

The source handler and informant can move forwards through the stages but they can also move backwards, or skip stages entirely. Three processes that are proposed to drive these movements are trust, similarity, and empathy. An increase in these processes moves the relationship to a more advanced stage and increases information disclosure. Conversely, a reduction in these processes can move a relationship backwards, or in the case of a severe violation to a party's expectations, can move the relationship to the termination stages.

Trust develops slowly across time. It is informed by multiple inputs as indicated by other aticles in this issue. Related research has emphaised the importance of integrity; an officer who fulfils their promises will foster trust and subsequently promote information disclosure. At the Initiating stage, beliefs regarding another's integrity will be relatively under-developed as there is

minimal information on which to evaluate the other's honesty, consistency or ethical values. At this stage, we might expect trust to be driven by factors such as a person's disposition, or general expectations/ beliefs about source handlers' ability to deliver on their promises if information is shared. These bases allow the relationship to develop, but do not necessarily result in rich information-sharing.

In contrast to trust, similarities (e.g., personal interests) are often used by source handlers as 'hooks' to build a connection. Similarity is influential from early on in a relationship – judgements regarding similarity within the very first interaction can guide later decisions about a relationship. Empathy is emphasised by source handlers as it creates a comfortable atmosphere and rapport. The cognitive, perspective-taking aspect of empathy is expected to contribute to intelligence gathering as a non-coercive tactic. Like similarity, empathy can also occur early on (e.g., source handler showing understanding of the informant's situation). Similarity and empathy could drive relationship progression in the early stages, but trust may take over as small trust exchanges are fulfilled and larger exchanges (or dependencies) develop. Although it takes time to develop, trust may be the most influential process in enabling information disclosure, by decreasing risk perception.

"Although it takes time to develop, trust may be the most influential process in enabling information disclosure.

PRACTICAL IMPLICATIONS

Applying the Staircase Model to source handler-informant relationships provides a framework to identify effective strategies for relationship development. It maps out processes likely to impact information disclosure at different points and allows us to consider how these processes interact and co-exist over time.

Andreea-Antonia Raducu is a CREST-funded PhD student at Lancaster University. Her research focuses on the lifecycle of the informant-source handler relationship and trust-building strategies.

Image credit: *rawpixel.com* | stock.adobe.com

STACEY CONCHIE & PAUL TAYLOR

TRUST SIGNALS

Beliefs about trustworthiness are central to security. In scenarios as diverse as military peacekeeping, vetting interviews, and bomb threat assessments, our appraisal of how much we trust another (the citizen, the candidate, the threat reporter) affects our view of risk and how we then act.

Most people see trust as a conscious judgement. We observe another's actions, apply meaning to these actions, and adjust our trust beliefs accordingly. We're aware that our conscious beliefs can be compounded by bias. In cross-cultural interactions, for example, trust beliefs can be confounded by conscious 'second-guessing' motivated by a desire not to antagonise or appear stereotypical. Ironically, the original first impression is often more accurate than the over-thought assessment.

The mention of first impressions hints at another facet of trust. Not all beliefs are formed consciously. Research shows that social interaction is governed by perceptions and beliefs that occur outside of conscious awareness. This includes trust beliefs, which are heavily shaped by subtle signals in our interpersonal behaviour. These signals are routinely detected, but they can fail to reach consciousness, or are overruled by conscious deliberation. This means that a person's natural capacity to interpret relevant social signals is often not utilised in their subsequent decisions and actions.

Several lines of research illustrate this point. Studies of peripheral vision show that people can detect threat without conscious awareness (i.e., fight or flight). DARPA capitalised on this evidence by building a military helmet that detects the brain's recognition of threat and brings this signal to the wearer's attention via haptic feedback. Equally, people's trust in another can be unknowingly shaped by altering the trustee's eye gaze. And, delaying the speed or type of nonverbal mimicry (movements that coincide with the timing and rhythm of a partner's movements) can render a viewer to distrust another.

NONVERBAL BEHAVIOURS

Inherent in these studies is the notion that automatic trust beliefs can be detected by subtle changes in the trustor's nonverbal behaviour. There are many parallels here with DARPA's work on physical risk detection through the measurement of brainwave activity. Instead, here we are interested in social risk detection (e.g., is the person trustworthy?), through the measurement of nonverbal behaviour. Data from two studies show what is possible.

Using methods from the film industry for capturing body movement—think of films like Avatar and Ted—we examined people's movements as they interviewed six citizens. Our citizens were actors who varied how much they cooperated with the interviewer and how much information they held. Some held no information of interest. Others held information that was either factual or false. Would the interviewer show different movement when interacting with the least trustworthy citizen (i.e., the person who had information, but was uncooperative)?

The results from sensors showed they do. More movement and more erratic movement betrayed our interviewers' unconscious lack of trust. Ostensibly, they were working harder to encourage an interaction with the citizen, even though they may not have realised it. Using movement data alone, it was possible to differentiate cooperative and non-cooperative citizens 22% better than guessing. This is the first hint that automatic trust beliefs, which develop rapidly, may be detectable through behaviour.

VERBAL BEHAVIOURS

Research has identified subtle trust signals in verbal behaviour too. This was perhaps first brought to life by Sandy Pentland's research, reported in Honest Signals, that found correlations in voice stripped of meaning (the sound retained but the vocalised words removed) can predict outcomes like team performance and negotiation outcome, which both depend on trust. Others have developed a 'trust dictionary' that claims to access subtle behavioural markers of a trustworthy speaker when applied to texts such as political speeches.

The recent PhD work of Steven Nicholson serves to illustrate what verbal signals can do. Across four studies he revealed that online dyads and online groups report greater trust when

People's trust in another can be unknowingly shaped by altering the trustee's eye gaze.

members are mimicked during early 'forming' stages and see more positive emotion language in later problem-solving tasks. However, this was reciprocal. While Steven could increase trust by injecting these behaviours, he also found that group members primed to believe their group was either high or low in trustworthiness would produce more positive emotion words and language mimicry.

His final study investigated online interactions in a virtual community focused on discussing credit card fraud (i.e. a criminal online group). When law enforcement intervened in the forum to disrupt activity, the group's language mimicry dropped significantly (in fact, Steven made this prediction before he was told when the disruption occurred). The intervention appeared to work, since the language change suggested a decrease in trust. But, after a period of two weeks, the mimicry recovered to the same level as before the intervention. It suggests the disruption impacted group trust for approximately two weeks, after which business returned to normal.

Trust signals may be a unique and useful form of intelligence for security, if we can find ways to harness them effectively and ethically.

Stacey Conchie is a professor in psychology at Lancaster University and Director of CREST.

Paul Taylor is professor of psychology at Lancaster University and the University of Twente.

Image credit: *shurkin_son* | stock.adobe.com

EMMA BARRETT

THE SCIENCE OF BETRAYAL: LESSONS FOR SECURITY PRACTITIONERS

An overview of research on betrayal from different disciplinary perspectives highlighting some important implications for defence, security, and policing contexts.

Trusted relationships are at the heart of security work: between staff working within security organisations, across organisational and national boundaries, and with members of the public who support security missions, such as covert human intelligence sources (CHIS). Without trust, information may not be shared, organisational relationships may be undermined, and operations may be derailed. Betrayal is a common reason for, and a common consequence of, a breakdown in interpersonal trust. It is also a common feature of intelligence work.

BETRAYAL IN DEFENCE, SECURITY AND POLICING

Most researchers agree that betrayal occurs when a trusted person, group, or organisation does something (or fails to do something) that causes someone to be harmed or wronged in some way.

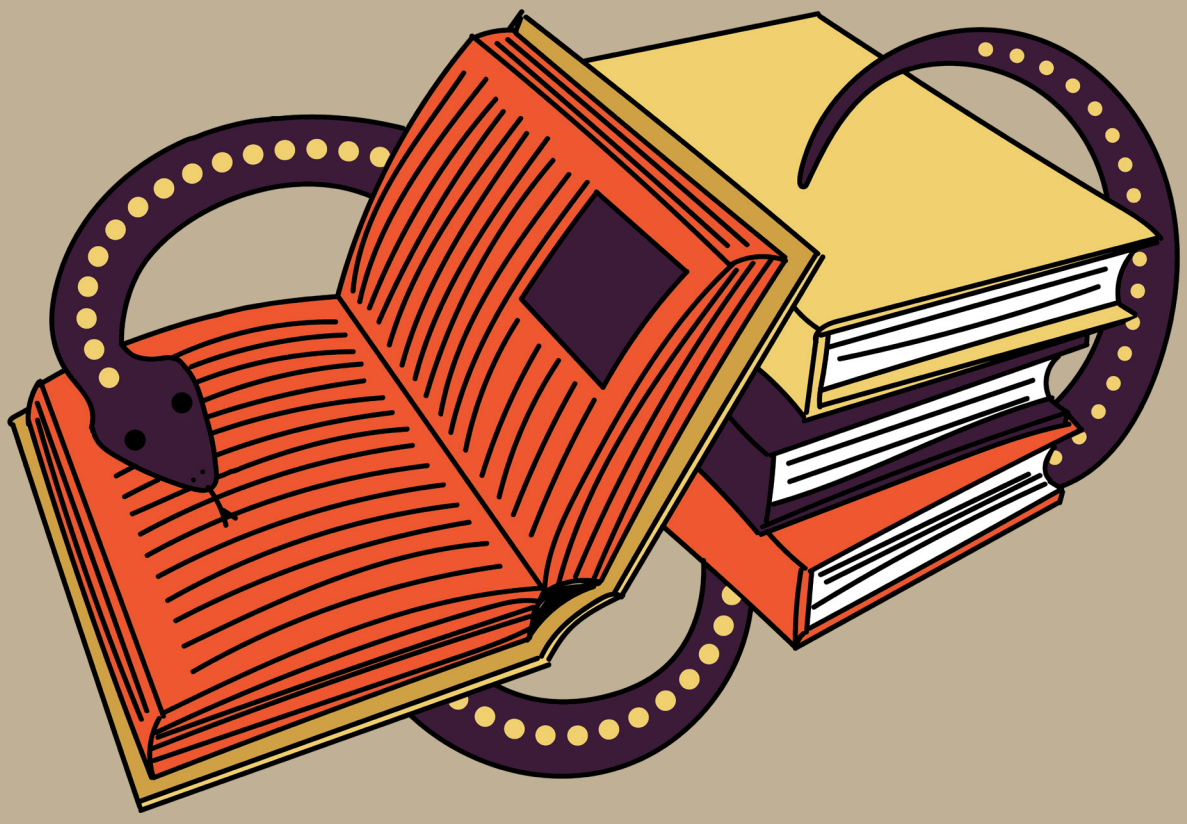
Betrayal crops up in many security contexts. It is a characteristic of human intelligence operations: the work of CHIS, spies, and undercover (UC) officers involves gaining or exploiting access to a group that is under investigation and betraying that group by passing on information that the group would prefer to keep secret so that security organisations can frustrate the group's activities.

Betrayal is also relevant to personnel security. Organisations that deal with sensitive information must guard against potential threats posed by 'insiders' – trusted members of an organisation who, like CHIS, betray their colleagues to a rival group. The most damaging insiders are viewed as traitors by their organisation or nation.

> **Betrayal is often viewed in a negative light. However, some betrayals can be seen as prosocial, such as whistleblowing and witness reporting.**

Criminal and terrorist behaviour often features betrayal. Many crimes involve a criminal building trust with their victims in order to betray it, usually for financial gain. Criminals and terrorists also deal with betrayal from each other, so will guard against informers or being ripped off.

Although betrayal is often viewed in a negative light, betrayals such as whistleblowing and witness reporting can be seen as prosocial. These acts involve bravery in stepping forward to report wrongdoing, and exposing illegal, immoral, or corrupt practices. But there is subjectivity and ambiguity here too: one person's heroic whistle-blower is another person's traitor, as in the cases of Edward Snowdon or Julian Assange.



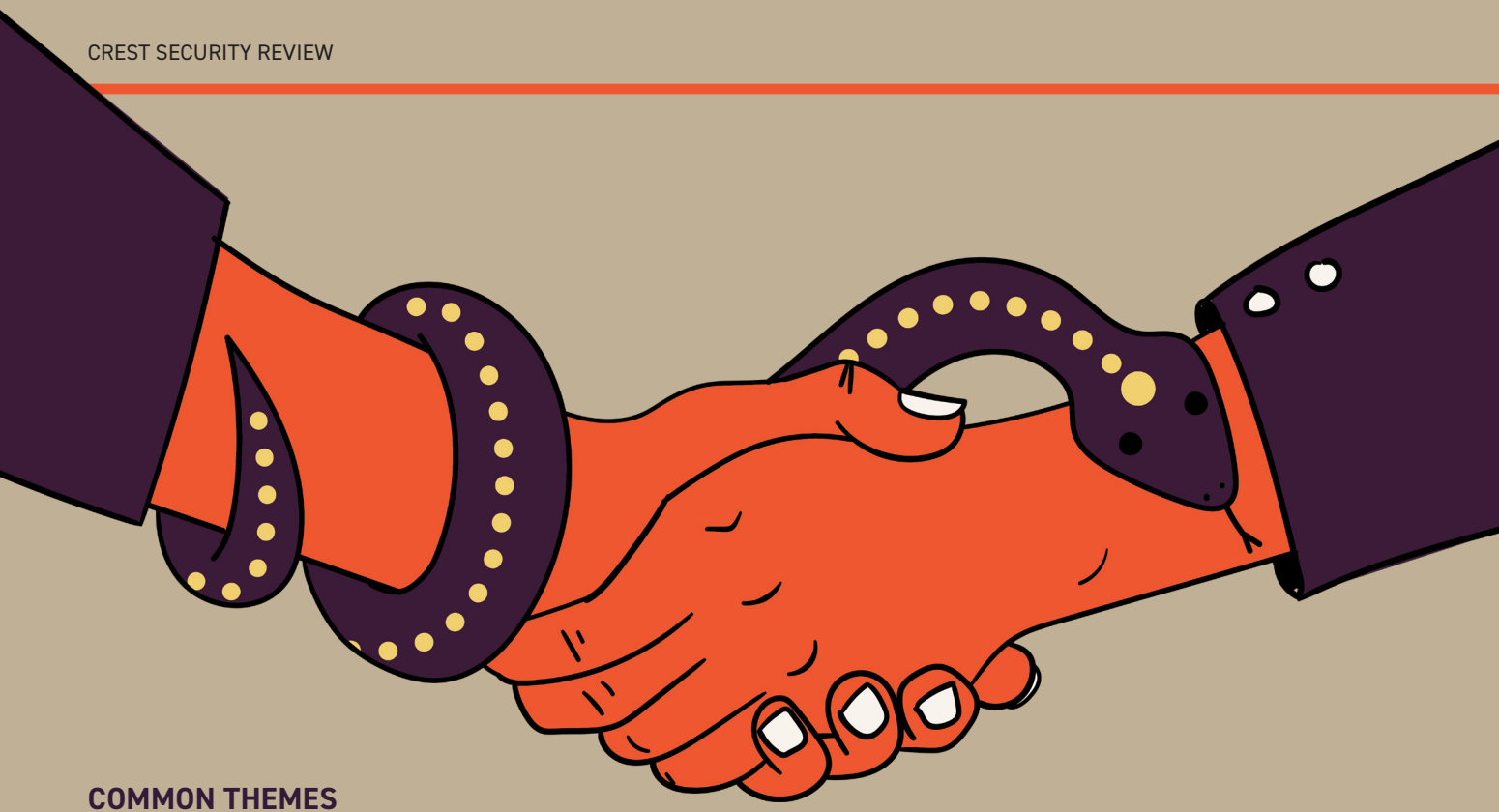
RESEARCH ON BETRAYAL

Although betrayal is rarely the sole focus, studies in security and policing contexts often touch on its nature and impact. For instance, researchers have studied the following:

- Motives and behaviours of CHIS, spies, and traitors (e.g., Akerstrom, 1986; Ben-Yhuda, 2001; Margalit, 2017), revealing the complex pathways that lead to someone betraying others.
- The wellbeing of UC officers, and the psychological impact of betraying others.
- The disengagement experiences of those who leave (betray) terrorist and criminal groups.
- Strategies used by criminals to avoid betrayal, for instance, through the way they communicate, how they establish and verify reputations in online criminal marketplaces, and how they detect and deal with informers.
- The shattering impact on victims of criminal betrayal, particularly in the case of romance frauds.
- How organisations can detect ongoing insider activity or keep out potential insiders through vetting.

Within organisational psychology there is a sizeable literature beyond work on insider threat that deals with themes of betrayal in workplace relationships, often related to breaches of the psychological contract between an employee and their employer. A subset of this research focuses on how organisations can repair trust with their employees after the psychological contract has been breached.

Beyond security contexts and looking across disciplines we find different perspectives on betrayal, and closely related concepts like loyalty, secrecy, deception, and of course trust. Several researchers focus on the experience of intimate betrayal, most often within romantic relationships, including when and how relationships are repaired. Another form of intimate betrayal, that of children by family members, caregivers, and institutions, has been explored in the context of childhood trauma and child development.



COMMON THEMES

A sense of betrayal can occur in a range of situations. In most security relevant contexts, betrayal is intentional. In some situations, however, a 'betrayer' may not realise that their actions will be perceived by the 'betrayed' as betrayal. For instance, a person may believe they have been betrayed but the alleged betrayer may argue that there was no expectation of loyalty or trust, and therefore the 'betrayal' was no more than a misunderstanding.

A second common theme is the complex relationship between, betrayal, trust and loyalty. Loyalty implies remaining true to a person, organisation, or cause, despite the existence of attractive alternatives. Being loyal implies you protect and defend the other party and its interests, even at your own expense. Trust does not necessarily require any of this, which means you can trust many different parties, but it is hard to be loyal to more than one. However, loyalty and trust both make someone vulnerable, and it is the exploitation of this vulnerability that is core to understanding the emotional responses to betrayal.

Regardless of context, betrayal elicits intense, often visceral, emotions, often many years after the act. A victim may experience:

- Anger towards the betrayer, and sometimes against others who allowed the betrayer to act.
- Humiliation and shame at having been fooled or manipulated by the betrayer, which can lead to self-directed anger at their gullibility.
- Sadness, disappointment and pain on realising that their trust in the betrayer was not met, or in reaction to a financial or other material loss.
- Confusion, loneliness and isolation as they try to make sense of how a betrayal occurred, and why the possibility of betrayal was not noticed or guarded against.

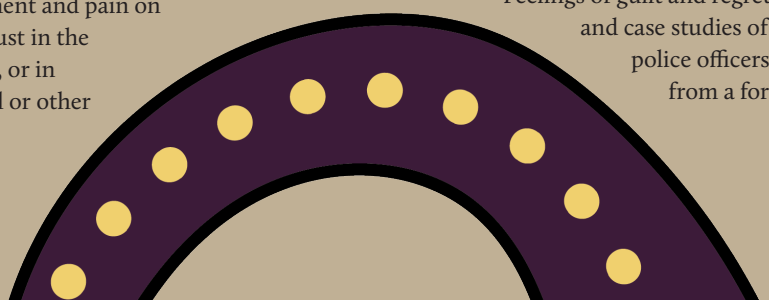
These emotions are felt across contexts, from romantic betrayals to organisational ones. And people can feel these emotions even when the betrayal was not directed at them. For example, retired CIA officer Jack Devine, talking about finding out that a colleague had volunteered to spy for Russia said:

> *I knew [Aldrich Ames] personally, went to his wedding. And his is one of those great agonies in life to know, personally, someone you would consider, ...a friend or at least be friendly with, that they betrayed their country. So that was the CIA case that was very disturbing to all of us...*

THE IMPACT OF BETRAYAL ON THE BETRAYER

Betraying another can generate intense emotions. Some may be positive: pride in exposing wrongdoing, or delight at manipulating others. But negative feelings are common, regardless of whether the betrayal is exposed. Betrayers may fear discovery and the consequences of betrayal: keeping a difficult, embarrassing, or traumatic secret is psychologically stressful. They may feel guilty about the impact on those they have betrayed, and shame at their actions, even when the betrayal is prosocial.

Feelings of guilt and regret regularly feature in research and case studies of the impact of UC work on police officers, as illustrated in this comment from a former UC officer:



A betrayal is a powerful signal that the victim and their needs have been devalued, setting up or reinforcing a power imbalance.

> *I've done nothing but spend every moment that I'm with them ensuring that they trust me... so that in the end I can use everything that they've said and done against them... when you think of yourself as a good person that kind of goes against that.*
>
> *Quoted in Coghlan, 2010*

One of the reasons for feelings of shame and guilt is the pervasive stigma around betrayal: positive synonyms for betrayal are rare but we have many negative words for someone who betrays: snake, snitch, tattletale, rat, grass. Contrast these with the positive sentiments that we have for people who don't betray: they keep secrets, they demonstrate loyalty, they can be trusted by their group.

It is not surprising that even prosocial betrayal can cause psychological stress for those who must betray as part of their job. Scant research has focused specifically on betrayers' coping mechanisms, but some strategies are evident in broader research and in case studies.

One common strategy is rationalisation and justification. Aimen Dean, an informer against Al Qaeda, explains that part of his coping strategy was the mantra "betrayal of the treacherous is loyalty in the eyes of God...", repeatedly justifying his betrayal as something that would please his God.

Another coping strategy is compartmentalisation. Kim Philby, unmasked as a Russian spy in the 1960s, wrote:

> *I have always operated on two levels, a personal level and a political one. When the two have come into conflict I have had to put politics first. The conflict can be very painful. I don't like deceiving people, especially friends, and contrary to what others think, I feel very badly about it.*
>
> *Quoted in Macintyre, 2014*

RESPONSES TO BETRAYAL

Although there are many varieties of betrayal, relatively fewer options exist for responding to it. Betrayal is often terminal for a relationship, but there may be ways of repairing the damage.

A betrayal is a powerful signal that the victim and their needs have been devalued, setting up or reinforcing a power imbalance. Some of the responses for betrayal are thus about rebalancing power. Acts of revenge are an attempt to deal with feelings of humiliation and anger, providing a sense of regaining control and getting even, although victims may continue to struggle with vengeful thoughts. An apology also seeks to rebalance power through a show of humility by the betrayer, though of course this will only work if the victim accepts this as a genuine apology.

Some victims ignore betrayal – something that psychologist Jennifer Freyd characterised as betrayal blindness. People often disregard or fail to look for signs of betrayal if they depend on the relationship for something important. In security contexts, employees and employers could be blind to potential insider behaviours because becoming aware of them disrupts workplace relationships. Or a handler might be blind to betrayal by their CHIS (e.g., signs of being a double agent) because they produce good intelligence.

Both sides also learn lessons from a betrayal experience. The victim may learn ways of coping or to be more wary of relationships. The betrayer may learn never to do it again, or how to get away with it.

KEY TAKEAWAYS

1. Betrayal has objective and subjective qualities. In defence contexts, acts of betrayal are often clear cut, as with insiders, spies, and informants. But it can also be a subjective judgement, and it is possible to betray someone or something without being aware that one has done so. For example, organisational change might be seen by some employees as a betrayal of the psychological contract, perhaps setting up a situation where an employee seeks vengeance. A handler may unwittingly do something that their CHIS views as a violation of trust and loyalty. When relationships start to go awry, it might help to consider whether perceptions of betrayal might be relevant. And when planning change, considering whether this might be viewed as a betrayal might help with your actions or your communications.
2. Betrayal has an emotional impact on betrayers as well as victims. Organisations that use UC officers and CHIS, or who want people to come forward to blow the whistle on bad behaviour, need to take account of the potential emotional impacts. Discussing negative feelings and developing coping strategies can help to limit the negative consequences of betrayal.

Emma Barrett is the Professor of Psychology, Security, and Trust at the University of Manchester.

ROSALIND H. SEARLE

TRUST = CONFIDENCE + VULNERABILITY
THE ROLE OF THE LEADER

Effective relationships are those that rely on trust. Trust has been described as the glue that sticks relationships together, or the oil that keeps them running smoothly. However, some relationships are more significant for trust than others, such as those with a senior or line manager.

VULNERABILITY IS TAXING

Trust involves two distinct facets: confidence in the other party, and a willingness to make oneself vulnerable. While a great deal of prior attention has focused on understanding the components of confidence, far less attention has been paid to the vulnerability involved in trusting the other party where there might be little means to control or monitor their behaviour. Feeling vulnerable diverts cognitive resources toward mitigating the perceived threat the other party poses. At its most extreme, where perceived risks outweigh the benefits, it can lead to the relationship being curtailed. Efforts to mitigate vulnerability raise the need for controls, which carry time and financial costs , but also divert effort from task performance into monitoring the other party's actions and compliance.

An indirect consequence of vulnerability is the introduction of additional stress and strain , which over time can further deplete the resources of the trusting party. This stress can introduce unintended errors into the individual's work, creating further costs and unintended security consequences for the organisation. Indeed, the trusting party may not be aware of the impacts of additional cognitive burdens on their decisions and actions – only realising after making a mistake.

> **Rule-breaking by leaders was associated with wide-ranging counterproductive work behaviours.**

THE CRITICAL ROLE OF LEADERS

Our CREST-funded study of trust in a high-security context showed that leaders play a critical role in trust, in part by shaping felt vulnerability. Leaders act as powerful role models who anchor others' behaviour. In this way, they influence how much confidence a person has in others' competence, their adherence to moral principles, and their care and respect for others' needs (i.e., the confidence facet of trust). The behaviours they promote also shape how much vulnerability a person experiences. We found that leaders who were immoral promoted vulnerability, reduced trust and increased security risks, while leaders who were moral mitigated these effects.

RULE-BREAKING AND MISCONDUCT

Our work showed that rule-breaking by leaders was associated with wide-ranging counterproductive work behaviours. Rule-breaking removed a leader's moral authority, allowing subordinates to perceive that they could do similar, creating the start of collective moral disengagement about rules and to whom they apply. This process undermined coherence

within the team, significantly reducing their capacity to contain wrongdoing through social sanction. In cases where team members were more pervasive and wide-ranging in their misconducts, vulnerability within the team increased and new stresses were created. For those not engaging in misconduct, they could either stay silent, exit the organisation or join in.

Collectively these deleterious processes reduce the performance of the team, replacing organisational interests with more self-serving goals. More critically, they re-shape local and organisational norms, creating a form of 'frog-boiling' as collective moral disengagement becomes normalised.

This effectively diminishes the means of social sanction and emboldens those engaged in misconduct and the leader in further self-serving antics. As feelings of vulnerability escalate within the team, a pernicious erosion of trust occurs. More concerned members start to quit, only to be replaced with self-serving individuals who are increasingly attracted to the team. In this way, the organisation can start to rot from within, with the means of self-correction diminishing rapidly. It is here that security risks are greatest.

...the organisation can start to rot from within, with the means of self-correction diminishing rapidly. It is here that security risks are greatest.

ETHICAL LEADERS

In contrast, ethical leaders offer a means to build and sustain teams and organisations that are resilient to security risks. These leaders are principled, honest and caring (thus building the confidence facet of trust) and operate by clear ethical standards, which they communicate to their followers. Ethical leaders discourage subordinates from regarding rules as things that are imposed on them as a means to control behaviour (i.e., to gain a reward or avoid punishment), but instead use these to enforce ethical standards. Ethical leaders encourage subordinates to model their behaviour in novel situations to determine for themselves what is right.

The efforts of an ethical leader diminish feelings of vulnerability as subordinates can understand the basis for their leader's decisions and actions, freeing them to concentrate on the task at hand, rather than being diverted to self-protection. They provide adherence to and development of effective systems, and challenge those that are not effective. These leaders build trust not only with their followers, but more widely outside of their team. It is therefore an important style of leadership offering important assurances to external stakeholders, that enhance the viability and resilience of the organisation, especially during times of crisis.

Rosalind H. Searle is a professor in human resource management and organisational psychology at the University of Glasgow.

Image credit: *Nattapol_Sritongcom* | stock.adobe.com

STEVEN LOCKEY

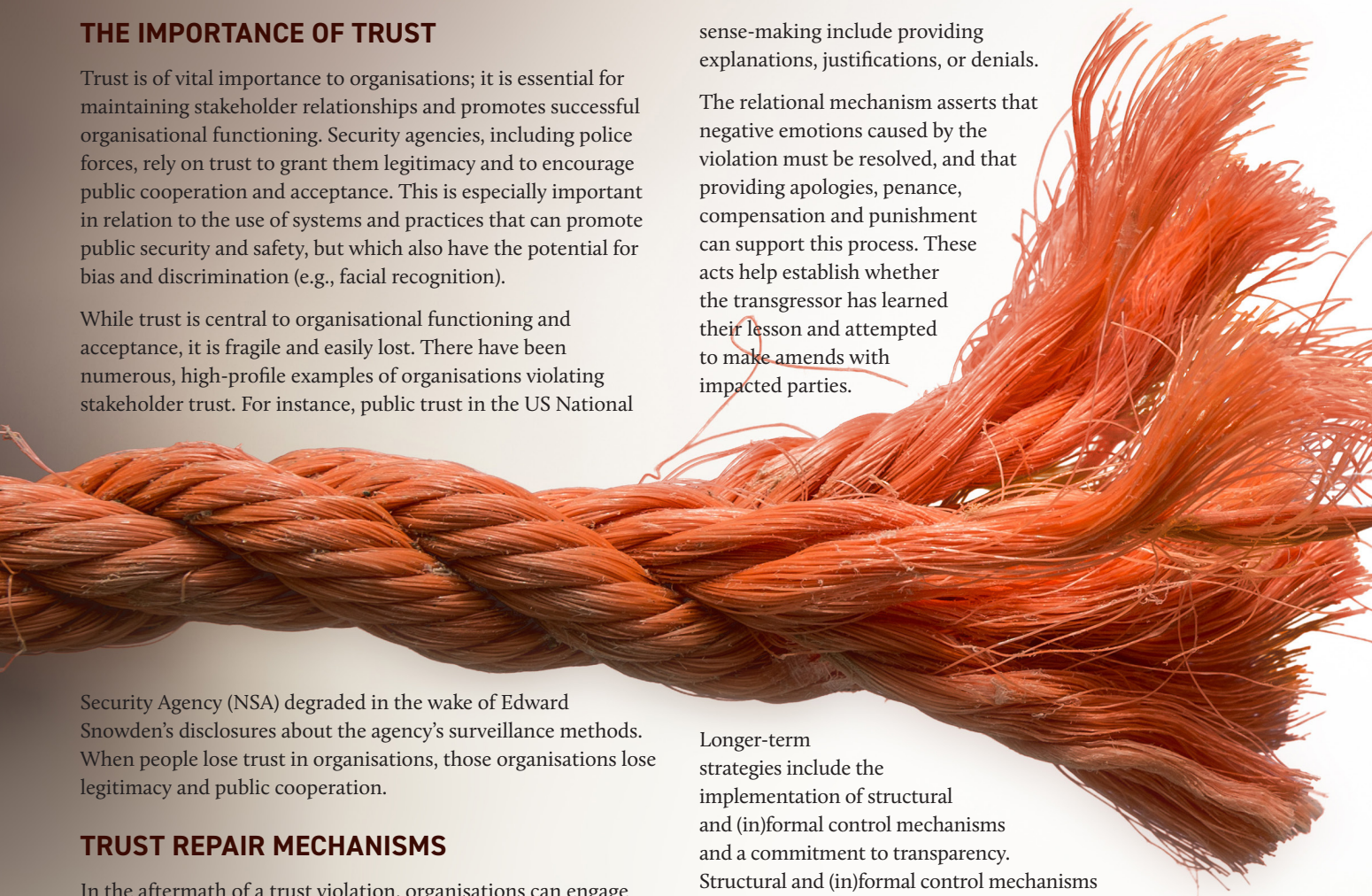
RECOVERING FROM FAILURE:
WHAT CAN SECURITY SERVICES DO TO REPAIR TRUST?

Trust is crucial for organisational effectiveness, but how can companies respond if they violate stakeholder trust? Steven Lockey draws on the scholarly organisational trust repair literature to provide answers to this question.

THE IMPORTANCE OF TRUST

Trust is of vital importance to organisations; it is essential for maintaining stakeholder relationships and promotes successful organisational functioning. Security agencies, including police forces, rely on trust to grant them legitimacy and to encourage public cooperation and acceptance. This is especially important in relation to the use of systems and practices that can promote public security and safety, but which also have the potential for bias and discrimination (e.g., facial recognition).

While trust is central to organisational functioning and acceptance, it is fragile and easily lost. There have been numerous, high-profile examples of organisations violating stakeholder trust. For instance, public trust in the US National Security Agency (NSA) degraded in the wake of Edward Snowden's disclosures about the agency's surveillance methods. When people lose trust in organisations, those organisations lose legitimacy and public cooperation.



TRUST REPAIR MECHANISMS

In the aftermath of a trust violation, organisations can engage in both short-term and longer-term strategies to repair trust. Short-term strategies can include sense-making and relational mechanisms. Sense-making assumes that stakeholders need to know what went wrong and why it happened for trust repair to take place. This mechanism focuses on providing wronged parties with information that enables them to overcome negative perceptions about an organisation. Specific strategies to enable sense-making include providing explanations, justifications, or denials.

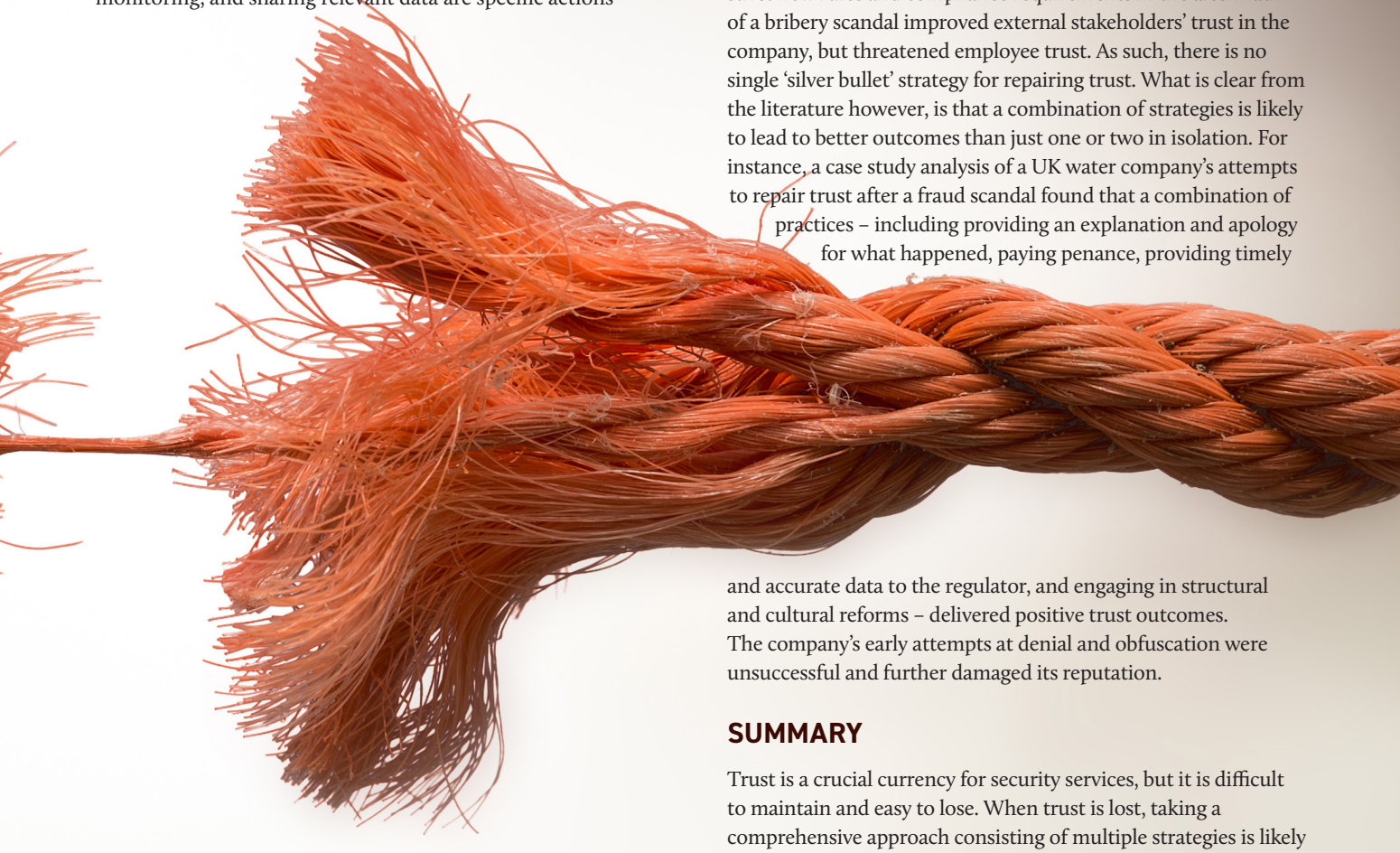
The relational mechanism asserts that negative emotions caused by the violation must be resolved, and that providing apologies, penance, compensation and punishment can support this process. These acts help establish whether the transgressor has learned their lesson and attempted to make amends with impacted parties.

Longer-term strategies include the implementation of structural and (in)formal control mechanisms and a commitment to transparency. Structural and (in)formal control mechanisms put in place rules or (in)formal controls that constrain the possibility of future transgressions and untrustworthy conduct. Specific strategies include implementing new policies, codes of conduct, incentives, sanctions, cultural reforms, and regulations. Changing formal structural and regulatory processes, and attempting to instigate cultural change are clearly time-consuming, costly, and difficult, but they are important in that

When trust is lost, taking a comprehensive approach consisting of multiple strategies is likely to produce better results than a piecemeal or reticent approach.

they demonstrate a substantive commitment to change. Returning to the Snowden NSA leaks, the US Government enforced a structural response by passing the USA Freedom Act in 2015 to limit the bulk collection of the telephone data of US citizens by the United States Intelligence Community (USIC).

The transparent reporting and sharing of information in the aftermath of a violation demonstrates that the transgressing organisation is behaving in a trustworthy manner. Conducting independent audits and reporting the results, allowing ongoing monitoring, and sharing relevant data are specific actions organisations can take in this regard. For instance, providing transparent access to police data has been proposed as a way to promote trust between the police and the community, particularly when a controversial incident occurs. Giving stakeholders access to statistics allows interested parties to determine how their local police force performs on salient outcomes. In turn, this can support them to make contextually accurate inferences, rather than assuming that a problem in one area is representative of all areas.



NO ONE-SIZE-FITS-ALL APPROACH

The mechanisms and strategies described previously can help organisations repair trust. However, that does not mean that repairing trust is easy. It is inherently complex, as intimated by the variety of cognitive, emotional, and structural processes underpinning the mechanisms. The complexity of trust repair is exacerbated by the fact that a variety of stakeholders have an interest in an organisation's activities, including employees, customers, suppliers, regulators, and the general public. These diverse stakeholders have different interests, power relations, and expectations about organisations and how they respond to trust failures. Indeed, trust repair efforts may enhance the trust of one stakeholder group but could further undermine the trust of other stakeholders. For example, Siemens' introduction of strict new rules and compliance requirements in the aftermath of a bribery scandal improved external stakeholders' trust in the company, but threatened employee trust. As such, there is no single 'silver bullet' strategy for repairing trust. What is clear from the literature however, is that a combination of strategies is likely to lead to better outcomes than just one or two in isolation. For instance, a case study analysis of a UK water company's attempts to repair trust after a fraud scandal found that a combination of practices – including providing an explanation and apology for what happened, paying penance, providing timely and accurate data to the regulator, and engaging in structural and cultural reforms – delivered positive trust outcomes. The company's early attempts at denial and obfuscation were unsuccessful and further damaged its reputation.

SUMMARY

Trust is a crucial currency for security services, but it is difficult to maintain and easy to lose. When trust is lost, taking a comprehensive approach consisting of multiple strategies is likely to produce better results than a piecemeal or reticent approach.

Dr Steven Lockey is a postdoctoral research fellow at The University of Queensland. His research interests include how organisations can repair trust after violations and trust in emerging technologies.

Image credit: *Romolo Tavani* | stock.adobe.com

ELLA GLIKSON

EMOTIONAL OVER-TRUST IN AI TECHNOLOGY

Our trust in AI technology is based on an evaluation of how we feel about it and how it performs. However, when we cannot evaluate its technological performance, this emotional-based trust can easily become a problematic over-trust.

We need a minimal level of trust to use any type of new technology. Some of this trust is based on rational thinking (e.g., the new technology's predicted reliability and usefulness), and some trust is grounded in emotion (e.g., linked to the extent we like the way the new technology is presented).

> **...anthropomorphic features such as facial features, human-like voice, or physical form significantly increase liking and trust.**

User Interface (UI) specialists use psychological principles to make technology easier to use and improve its attractiveness and likeability. One of the most popular ways they can make AI more likeable is to emphasise different types of anthropomorphic features. Empirical research consistently demonstrates that factors such as facial features, human-like voice, or physical form significantly increase liking and trust.

When these human-like features (such as AI's immediate responsive behaviours to user movement or words) also signal a technological ability to perform the required task, the two elements of trust (liking and rationality) are aligned. However, what happens when the presence of human-like features overshadow the rational evaluation of the technological ability, and why is this important?

EMOTIONAL VS RATIONAL THINKING

Human-like cues lead to high expectations about AI's technological performance. Research shows that the more technology is presented as a living organism, the more we like it and believe in its capabilities and moral values. For instance, giving an automated car a human name can increase our liking and trust, and lead to assumptions about the car's high performance and reliability. However, the level of technological attractiveness is based on the employment of psychological principles (e.g., similarity to a living thing or to a specific user) and has little to do with its algorithmic functions.

> **AI's human-like behaviours likely impact our emotions more profoundly than our rational thinking.**

AI's human-like behaviours likely impact our emotions more profoundly than our rational thinking. In several studies, researchers found that people tend to trust anthropomorphic robots, even when their low-performance ability was evident. Although these studies were performed in labs, where the actual implication of robotic performance is questionable, they raise an important query regarding the relative power of the emotional basis of trust in technology. The more complex the outcomes of algorithmic performance, the more difficult it is to correctly evaluate reliability, and thus the role emotions play in the evaluation become more significant.

PREVENTING OVER-TRUST

The disassociation between technology's likeability and its actual reliability and performance can be highly problematic, resulting in over-trust. Over-trust relates to a situation in which high trust in unreliable technology would lead to misuse, which may cause a breach of safety or other undesirable outcomes. Based on people's tendency to resist change, research tends to focus on ways to improve trust and facilitate the use of new technologies. Therefore, a lot of effort is made to understand how to improve the likeability of bots and robots and make them more integrated into organisations and everyday life.

Although this effort can result in an eagerness to use new technologies, there is a growing need to better understand how to balance the positive emotions evoked by technology's external features and the need for a rational evaluation of technology's reliability and performance.

Aiming to find new ways in which our (manipulated) emotional reactions will not lead to over-trust in technology that is biased, erroneous, or just not yet ready to perfectly perform the task at hand, we need to put more effort in demonstrating this phenomenon in lab and on-line experiments, as well as communicating the possible dangers to those with responsibility for purchasing new technology, and potentially also those with responsibility for regulating its use.

Ella Glikson is an assistant professor at the Graduate School of Business Administration in Bar Ilan University.

Image credit: *Brazhyk* | stock.adobe.com

PAUL TAYLOR

TRUSTING A CENTRE MODEL

Founding director of CREST, Paul Taylor, discusses the long-term benefits of trusting a centre model for research.

To misquote Seba Smith (1840), "There are more ways than one to skin a cat, so there are more ways than one to spend research funding." On concluding my tenure as CREST's director, I was asked by the research council who administered our funding to reflect on the value of a centre model. This is a question that should concern all research teams and government funders in our community, so at the risk of being self-indulgent: What does CREST do that cannot be achieved by contract funding or an in-house research team?

> **Some of their biggest successes are projects that resonate not today but as time passes.**

RISK MITIGATION

For those responsible for stopping today's threats, it is both reasonable and understandable to want agile, rapid research that helps today. But dealing with today means not preparing for tomorrow. Centres spread their bets across today and tomorrow. Some of their biggest successes are projects that resonate not today but as time passes. Who could have anticipated, back in 2016 when she started her CREST-funded PhD, that Christina Winter's analysis of vetting contexts and how to get the most out of video-enabled interviews would have been so prescient?

KNOWLEDGE QUALITY

Most knowledge in academia is never published—it's in a file drawer under 'failed attempts' or 'stuff everyone knows.' Centres have access to this file drawer and can expose it fully, giving users a truly balanced answer. They can also promote knowledge quality. Centres can applaud null findings and ensure their science is open in ways that are challenging for commercial models who depend on taking the next step.

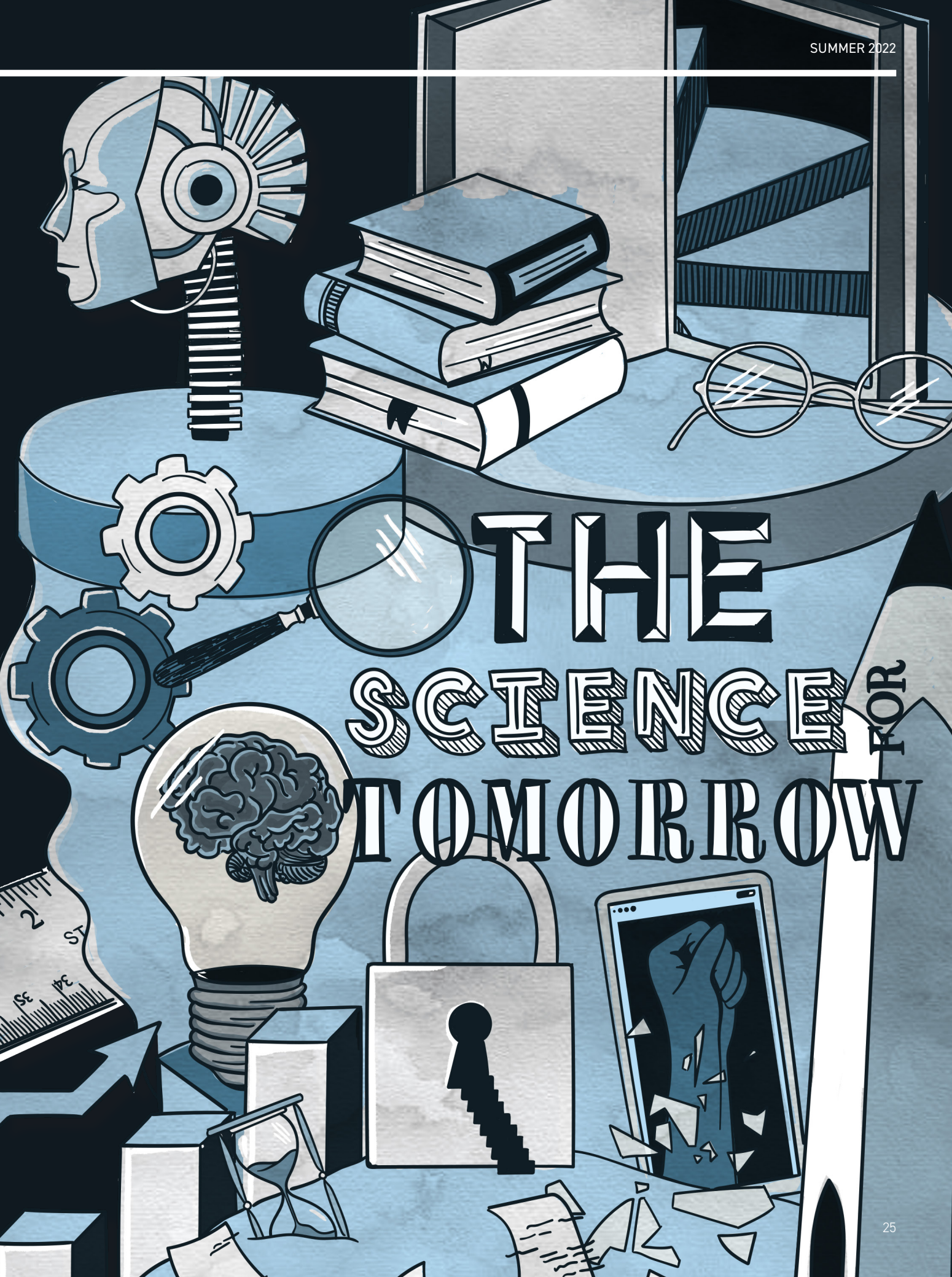
ECONOMIES OF SCALE

Over its first five years, CREST researchers secured a further £23m of follow-on funding to continue projects relevant to the needs of the UK's security and intelligence community. That's more than a three-fold return. Centres can start balls rolling in ways that single projects can't. Leading one of CREST's short-term, commissioned projects, Karen Douglas's review of the psychology of conspiracy theories was an important flag in the ground of a now burgeoning area. Elsewhere, the rapid emergence of cross-cultural interviewing research was driven forward by a focus across several of CREST's projects.

IMPACT

By far the biggest difference a centre can bring is impact. The visibility of a centre becomes a focal point not solely for specialists but for a wider community with an appetite to learn and apply best practice. An independent review of CREST's impact revealed examples of evidence being used to support operations, training, and tradecraft (Edwards, 2020). Yet, reviews are restricted to what is known, and the reach of a centre into the culture and thinking of an organisation is far greater than what is measured by a Likert scale.

Paul Taylor is professor of psychology at Lancaster University and the University of Twente. At Lancaster, he founded and directed CREST from 2015 to 2021. In May 2021, Paul became the first UK Chief Scientific Advisor for Policing.


THE
SCIENCE
FOR
TOMORROW

CALVIN BURNS

A-Z OF TRUST

Trust has become an essential concept in security research. This A to Z provides an overview of how trust can be conceptualised, measured and influential in shaping behaviour.

ASYMMETRY

Trust is harder to build than it is to lose. Slovic (1993) demonstrated trust asymmetry by showing that negative events have a much more significant impact on trust than positive events.

BEHAVIOUR

Some researchers conceptualise trust as choice behaviour and study it using prisoner dilemma games. Most researchers, though, distinguish between 'trust' and 'trusting behaviour' by clarifying that trusting behaviour involves assuming risk or 'risk-taking in a relationship' with another person, whereas 'trust' involves a willingness to assume risk.

COGNITION-BASED TRUST

Trust has a cognitive basis in that it is based in part on perceptions of trustworthiness and other factors, like value similarity or group membership, which give an individual good reason to be willing to take a risk with another person.

DISCLOSURE OF SENSITIVE PERSONAL INFORMATION

Disclosure of sensitive personal information is a behaviour that may result from trust, developed for example, during an investigative interview. Current CREST research is investigating the role of trust in the disclosure of sensitive personal information under different conditions.

EMOTION

Emotion is part of Affect-based Trust, proposed by McAllister (1995). Affect-based Trust is thought to develop from some level of Cognition-based Trust and can develop into deeper forms of Relational Trust.

FACTORS OF PERCEIVED TRUSTWORTHINESS

Trust beliefs or perceptions of trustworthiness can be based on many factors. Mayer et al. (1995) proposed Ability, Benevolence, and Integrity as three factors that can account for most of the variability in perceptions of trustworthiness.

GAMES

Trust games (or prisoner dilemma games) have been used to study trusting behaviour and how people make decisions about trust. They usually involve two stages. During the first stage, Player 1 chooses between a guaranteed outcome or trusting Player 2. If Player 2 is trusted, then Player 2 decides whether to reciprocate or betray Player 1's initial act of trust.

HISTORY-BASED TRUST

Models of History-based Trust assume that trust develops as a function of cumulative interaction. Instances of past behaviour are used to make decisions about another individual's trustworthiness and likely future behaviour.

IMPLICIT ATTITUDES ABOUT TRUST

Attitudes about trust can be activated automatically and measured implicitly, usually by reaction times. Automatically activated attitudes about trust are thought to influence behaviours that individuals do not try to control consciously and thus may be indicative of deeper forms of trust.

JOINING GROUPS

Research by Morrison (2016) suggests that trust may be more important than ideology when deciding which side to join when terrorist groups split. Trust that results from information about someone's membership in a social or organisational group is called Category-based Trust.

KNOWLEDGE-BASED TRUST

Some researchers differentiate between 'Calculus-based Trust' and 'Knowledge-based Trust'. Real trust (as involving a willingness to accept vulnerability) is thought to start with Knowledge-based Trust or positive confidence based on prior predictability.

LEADERSHIP

Most leadership researchers recognise trust as an important concept. Transformational and charismatic leadership models suggest that leaders build trust in their followers. Trust is also important in leader-member exchange theory as higher levels of trust are associated with higher quality exchange relationships and in-group membership.

MISTRUST

Mistrust, or Distrust, is characterised by a lack of trust. Some researchers conceptualise trust and mistrust at opposite ends of the same continuum. Other researchers conceptualise trust and distrust as separate constructs; they consider distrust to be based partly on confident negative expectations regarding another's conduct.

NO CHOICE BUT TO TRUST YOU

If someone says, 'I have no choice but to trust you,' it is not real trust because there is not a willingness to be vulnerable. Situations like this have been conceptualised as Calculus-based Trust, in which the person taking the risk has suspicions, but the benefits outweigh the costs.

ORGANISATIONAL TRUST

Three broad types of Organisational Trust appear in the literature: 1) Trust within organisations (e.g., between employees or co-workers, or between workers and management), 2) Trust between organisations and their customers (e.g., for marketing purposes), and 3) Trust between organisations.

PERSONALITY TRAIT

Some researchers conceptualise trust as a personality trait which can be measured by propensity or predisposition to trust questionnaire scales.

QUALITATIVE DEGREES OF TRUST

Dietz and Den Hartog (2006) identified five Qualitative Degrees of Trust in the literature (Deterrence-based, Calculus-based, Knowledge-based, Relational-based, and Identification-based). These qualitative degrees reflect trust in different sources but also different types of trust experience.

REPAIRING TRUST

Trust is fragile and, when broken, has serious consequences for individuals and organisations. Trust can be repaired. The two dominant trust repair strategies are short-term (e.g., excuses, apologies, denials) and long-term (e.g., remaining silent, structural rearrangements).

SWIFT TRUST

Swift Trust was proposed by Meyerson et al. (1996) to explain trusting behaviour by members of new project teams or people working in temporary organisational structures who had no past working relationships with each other. Some researchers have suggested that it is not a form of trust but a trust substitute or risk management strategy.

TCC MODEL

Some researchers have argued that trust is strongly related to risk perception, while others have argued that the two are weakly related. The Trust Confidence and Cooperation (TCC) was proposed by Earle and Siegrist (2008) to explain the relationship between trust and risk perception, mainly in the context of risk communication.

UNCERTAINTY

The world is an uncertain place. If we could predict the future with perfect certainty, trust would not be needed. In that sense, trust and uncertainty are opposite sides of the same coin.

VULNERABLE

Most researchers today conceptualise trust as a willingness to be vulnerable or take a risk. (See Ros Searle's piece)

WHY DO PEOPLE TRUST?

Some researchers have proposed that we developed trust as a heuristic or cognitive shortcut to help us cope with uncertainty and risk. Conducting systematic and detailed evaluations of every situation we face would overwhelm our cognitive capacities. Trust, therefore, allows us to get on with life and can lead to decreased transaction costs.

X-CULTURAL (OR CROSS-CULTURAL)

Cross-cultural research indicates that cultural values may influence how trust develops in different national groups.

YOUR EXPERIENCES

Your experiences of trusting another person will shape and reinforce your trust beliefs about that person. Over time, your trust experiences can influence your general propensity to trust other people.

ZAND

Trust was first proposed to be context-specific by Zand (1972). He suggested that it is possible to trust a person in one situation but not another. An example of this is 'Safety-specific Trust.'

Calvin Burns is a senior lecturer in occupational psychology at the University of Greenwich.

OLIVIA BROWN

RIGHT-WING EXTREMISM ONLINE:
CAN WE USE DIGITAL DATA TO MEASURE RISK?

The internet plays an important role in the rising threat of right-wing terrorism. Olivia Brown and colleagues have combined psychology and computational science methods to identify whether online behaviour can be used to infer the risk of offline action.

BACKGROUND

The threat of right-wing extremism is growing globally, with statistics showing a 320% increase in right-wing terrorist offences in the past six years. Evidence suggests the internet is playing a key role in this growth, with online forums and social networking sites providing the opportunity for individuals to share ideas, recruit new members, as well as offer a medium through which to acquire ideology and plan attacks (Scrivens, Gill, and Conway, 2020). This can be illustrated in recent high-profile incidents such as the Christchurch terrorist attack and Pittsburgh Synagogue Shooting, in which the perpetrators posted about their intentions online in the weeks and moments preceding their violent attacks.

In an increasingly digital world, the role of the internet in the planning and execution of terrorism presents law enforcement with an opportunity to build technological tools to assess online communications and detect risk. Supported by research in social psychology, there is a growing consensus that digital data may indicate when and how interactions online might lead to right-wing extremist violence offline. However, the volume of extremist content makes it challenging to identify which individuals pose a risk to public safety. The challenge of identifying these 'needles in the haystack' has been exacerbated by the pandemic, in which we have witnessed an unprecedented rise in extreme-right wing content, with members of the far-right exploiting anti-vaccine and anti-authority sentiment.

RESEARCH

With right-wing extremist content on the rise across mainstream and dark-web platforms, questions remain as to whether there are specific markers of online behaviour that can be used to infer risk. Our research begins to address this question by modelling the online interactions of right-wing extremists across three far-right platforms. Existing methods have tended to focus on large scale quantitative analysis of entire platforms, identifying patterns of posting and indicators of extreme content. While this provides an overview of the far-right online context, it cannot offer any indication as to how to identify users who may be at risk of committing a violent offence. Unique to our approach is the inclusion of data from individuals who have been convicted of a terrorism-related offence and those who have not. By using conviction as an independent variable and tracing our digital data back to specific individuals, our data presents a unique opportunity to develop insight on risk.

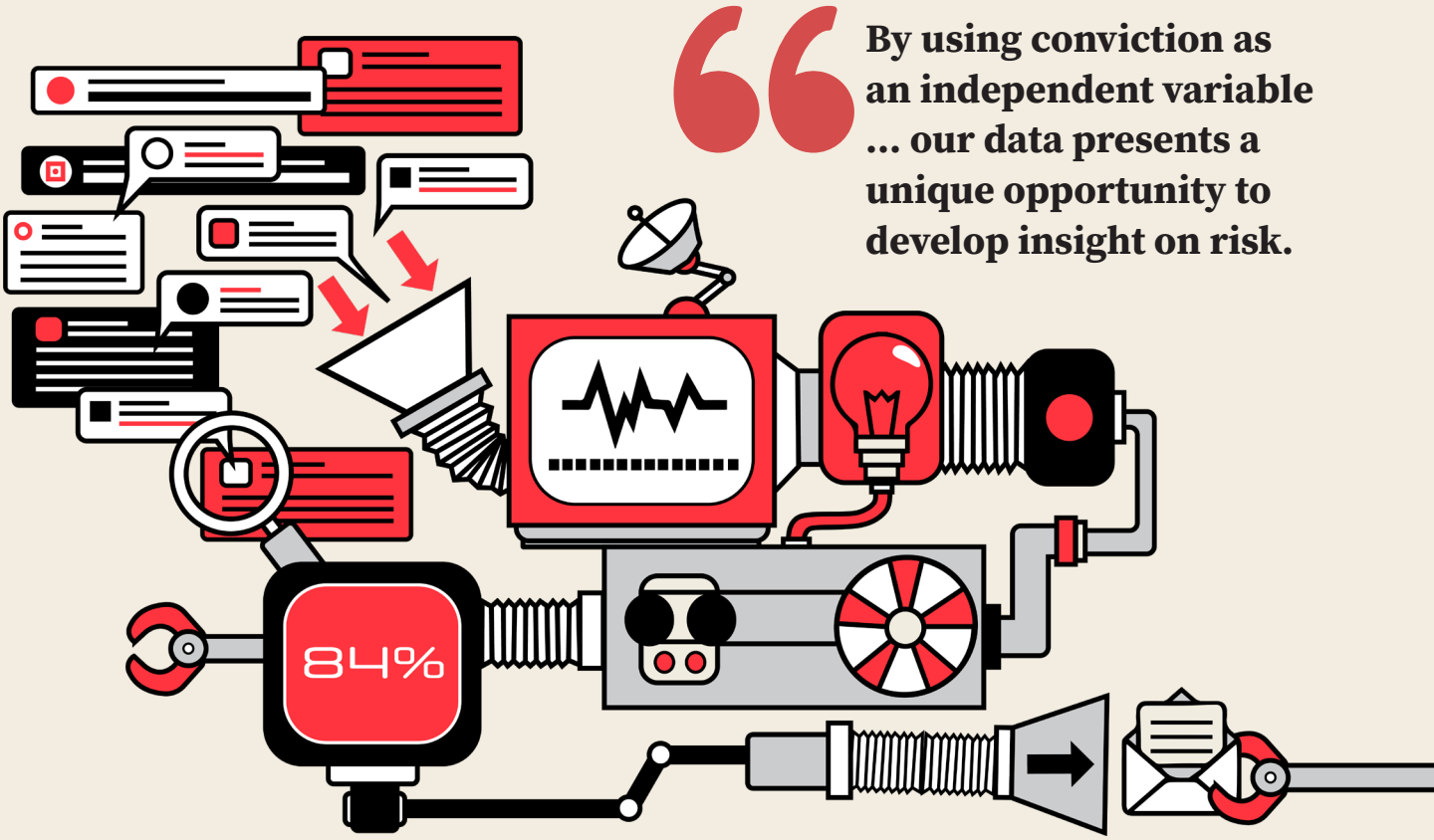
To compare convicted and non-convicted right-wing extremists, we obtained a sample of online postings and metadata that could be matched to individuals from their online aliases. All data obtained were publicly available and identified through open-source intelligence. We adopted strict inclusion and exclusion criteria to ensure that individuals (either convicted or not) were correctly matched to their online aliases. Our sample included 180,000 posts across three far-right forums (Gab, Discord, and Iron March) from 26 convicted and 54 non-convicted right-wing extremists.

We adopted a novel methodological approach to our analysis by combining qualitative and quantitative tools. First, a qualitative content analysis was conducted on 28,000 posts from eight convicted and eight non-convicted users. The qualitative analysis helped establish an in-depth understanding of the context of the research and began comparing users according to their conviction status. Notably, the results from the qualitative analysis were then used to inform our quantitative analysis. We adopted a computational approach to the quantitative analysis, in which we ran topic models to acquire features that could be used in a machine learning algorithm to predict conviction status based on post content.

WHAT WE FOUND AND WHAT IT MEANS

In the qualitative content analysis we identified 9 higher-order categories representative of the data:

1. Hateful content
2. Group Formation
3. Information Sharing
4. Intra-Group Debate
5. Operational and Security focused content
6. Threats of Violence
7. Offline Action
8. Weapons
9. Inciting Violence



By using conviction as an independent variable ... our data presents a unique opportunity to develop insight on risk.

For example, Group Formation represented content focused on forming groups, recruitment and connecting different online users together:

> *"Hi mate, I'm one of the main organisers with [REDACTED] in the UK. Who's currently in charge of [REDACTED]? I want to establish contact. Thanks"*

When comparing the two groups (convicted and non-convicted right-wing extremists), we found significant differences in each category of posts apart from Inciting Violence. Convicted users posted more of each category of content, apart from Intra-group Debate. Interestingly, this finding demonstrates that the convicted users were much more likely to be participating in direct action – whether that be through sharing files and websites (Sharing Information), discussing ways to avoid detection by the authorities (Operational and Security), working to recruit and connect others (Group Formation) or engaging in offline activities (Offline Action). Non-convicted users, on the other hand, were much more likely to discuss ideological positions, strategise about the movement going forward and debate historical events (Intra-Group Debate).

Next, we conducted topic modelling on the full dataset of 180,000 posts. We iterated through the topic models (ranging from 3-10 topics) to identify topics that were coherent with the qualitative analysis. Four topics were retained for further analysis - Hateful Content, Ideological Debate, Violence, and Intra-Group Connections. In the next step, we ran machine learning models to predict the conviction status of users (convicted versus non-convicted) using the degree to which their posts fit the topics (i.e., the probability distribution of each topic for each post) as features in the model. Our model currently demonstrates 84% accuracy when detecting whether a post belongs to a convicted or non-convicted user. When accounting for our unbalanced dataset (i.e., we have more posts from non-convicted than convicted users), the accuracy reduces to 65%. Notably, the findings demonstrate very high accuracy when detecting whether someone is non-convicted – 92%. This would suggest that our findings could be used to reduce the volume of digital data that requires close monitoring by the authorities by highlighting posts indicative of reduced risk (i.e., posts by individuals who, while expressing extremist views, have not engaged in activities that meet the threshold of illegal action).

CONCLUSION

With the volume and accessibility of extremist content online increasing, government agencies must continue to grapple with the challenge of identifying individuals who might be a risk to public safety. The methodology and findings presented here demonstrate promise and suggest cautious optimism in developing technological tools to narrow down the number of individuals in need of close monitoring online. We intend to continue building on this research and work towards developing algorithms that can identify 'bigger needles' and create 'smaller haystacks'.

Dr Olivia Brown is a lecturer at the University of Bath. She is interested in how intra- and inter-group processes influence individual and group behaviour.

BEN LEE

WHAT IS SIEGE CULTURE?

Although a fringe set of beliefs, Siege Culture has underpinned many of the recent counter terrorism cases linked to the extreme-right in the UK.

Siege Culture is the most extreme interpretation of fascism and national socialism yet seen. Siege Culture supporters have an anti-democratic, anti-enlightenment, racist and white supremacist worldview. They believe they are Aryans, a specific and superior group at the top of a racial hierarchy. They are hostile towards non-whites, non-heterosexuals, Jews, and government. Siege Culture includes the idea of 'The System', a conspiracy of the government, Jews, capitalists, and all other forces acting against Aryan interests.

Within Siege Culture, fascism is treated as a higher truth and a natural state in which Aryans will dominate all others. As a result, Siege Culture believes that any softening of their message to increase their appeal is impossible, and that any form of politics or compromise is inherently flawed. Siege Culture is critical of other right-wing actors who are seen as being insufficiently committed.

Siege Culture argues that societies are in a period of involution: a period of decay caused by weakness. The eventual collapse of society and destruction of The System is considered inevitable. Collapse is a necessary precursor to the rise of the organic state and a return to the natural hierarchy.

ACCELERATIONISTS

Accelerationism refers to a violent strategy in which terrorism is used to hasten societal collapse by provoking reactions from authorities and exacerbating existing social tensions. Although it did not originate with Siege Culture, the term has come to be closely associated with the space to the extent that Siege Culture inspired groups are often referred to as accelerationists.

> **Siege Culture is the most extreme interpretation of fascism and national socialism yet seen.**

OCCULTISM

Although cultic influences have been a persistent feature on the fringes of the extreme-right, from 2016 onwards occultism has played a greater role in Siege Culture. In some cases, this has taken the form of Christian Identity, Esoteric Hitlerism, and other beliefs that align heavily with racism. Since 2017, Left Hand Path Satanism including the groups Order of the Nine Angles and Tempel ov Blood, have also featured in Siege Culture. The incorporation of these ideas has been divisive and caused splits within Siege Culture.

AESTHETICS

Performance is a key aspect of Siege Culture. Activists linked to groups and brands are conscious of how they present themselves. Militancy, hypermasculinity, firearms, and neo-Nazi symbols are key aspects of online and (on rare occasions) public performance. Several key aesthetics have emerged from Siege Culture, most influential has been the work of Canadian propagandist Dark Foreigner.

> **Siege Culture is not a single ideology with a uniform set of beliefs. There is no acknowledged leader or single dominant personality.**

ORGANISATION

Siege Culture is not a single ideology with a uniform set of beliefs. There is no acknowledged leader or single dominant personality. The centre of the subculture is online. At various times this has included some key web forums (Iron March, Fascist Forge) and websites (Siege Culture, Noose, American Futurist). Siege Culture also persists on Telegram and other encrypted applications as well as less moderated platforms such as Odysee and Internet Archive. Some activists have founded small groups (such as Sonnenkrieg Division and Feuerkrieg Division). Online organising undoubtedly contributes to the lack of a uniform ideology and a strong transnational perspective. The decentralised nature of Siege Culture has left it vulnerable to ideological drift and introspection, including rifts caused by the influence of occultism.

Image: Exrtract from an Atomwaffen Fission pamphlet

VIOLENCE AND OFFENDING

The relationship between Siege Culture and violence is complex. Militancy is a key element of how Siege Culture presents itself. However, to date, successful and clearly identifiable right-wing terrorist attacks associated with Siege Culture have been rare. Plotting activity and actual violence have been far outstripped by online rhetoric and overall presentation.

Siege Culture has however contributed to the large upsurge in right-wing terrorism offending in the UK. The proscription of National Action in 2016 was a watershed moment and the numbers of right-wing terrorist offenders in prison, many convicted of membership offences, began to rise from 2017 onwards. Members of successor organisations, such as Sonnenkrieg Divison, have been convicted of terrorism offences including encouraging terrorism.

Despite proscription, the online subculture and networks that underpin Siege Culture remain persistent. Militancy and 'edginess' are a core part of the scene's aesthetic. Recent publications have sought to place renewed emphasis on direct action. It is currently not possible to say if these efforts have been successful at encouraging more violence. To read more of Ben's work on Siege Culture and accelerationism in the UK please see our CREST guide at *www.crestresearch.ac.uk/resources/siege-culture-and-accelerationism-in-the-uk*

Dr Benjamin Lee is a senior research associate at the Centre for the Study of Terrorism and Political Violence at the University of St Andrews where his research work is funded by CREST.

READ MORE

Read more about some of the research that our contributors mention in their articles. We've flagged up those that are open access and given links to online versions where they are available. On page 35 we also include articles relating to Trust that appear in our previous issues. For full references and citations please visit the online version at *crestresearch.ac.uk/magazine/trust*

BARRETT: LESSONS FOR SECURITY PRACTITIONERS FROM THE SCIENCE OF BETRAYAL

Åkerström, M. (1986). Outcasts in prison: The cases of informers and sex offenders. *Deviant Behavior,* 7(1), 1–12. *https://bit.ly/3agh5Ef*

Ben-Yehuda, N. (2001). Betrayal and Treason: Violations of Trust and Loyalty (1st ed.). Routledge. *https://doi.org/10.4324/9780429502071*

Crombag, H., Rassin, E., & Horselenberg, R. (2003). On vengeance. *Psychology, Crime & Law,* 9(4), 333–344. *https://doi.org/10.1080/1068316031000068647*

Curran, L. S. (2021). An exploration of well-being in former covert and undercover police officers. *Journal of Police and Criminal Psychology*, 36(2), 256-267. *https://bit.ly/3PbO8b9*

Fitness, J. (2001). Betrayal, rejection, revenge, and forgiveness: An interpersonal script approach. *Interpersonal Rejection*, 73-103. *https://bit.ly/3AA9FWU*

Freyd, J. J. (1997). II. Violations of power, adaptive blindness and betrayal trauma theory. *Feminism & Psychology,* 7(1), 22-32. *https://doi.org/10.1177/0959353597071004*

Grobbink, L.H., Derksen, J.J.L., & van Marle, H.J.C. (2015). Revenge: An Analysis of Its Psychological Underpinnings. *International Journal of Offender Therapy and Comparative Criminology,* 59(8), 892-907. *https://bit.ly/3ImqCGb*

Koehler, J. J., & Gershoff, A. D. (2003). Betrayal aversion: When agents of protection become agents of harm. *Organizational Behavior and Human Decision Processes*, 90(2), 244-261. *https://bit.ly/3alPWQa*

Macintyre, B. (2014). A spy among friends: Kim Philby and the great betrayal. London: Bloomsbury *https://bit.ly/3AALCHi*

Macleod, A. D. (1995). Undercover policing: A psychiatrist's perspective. *International Journal of Law and Psychiatry*, 18(2), 239–247. *https://doi.org/10.1016/0160-2527(95)00009-7*

Margalit, A. (2017). *On Betrayal.* Harvard University Press. *https://bit.ly/3n1oiwP*

Rachman, S. (2010). Betrayal: A psychological analysis. *Behaviour Research and Therapy*, 48(4), 304-311. *https://doi.org/10.1016/j.brat.2009.12.002*

BROWN: RIGHT-WING EXTREMISM ONLINE: CAN WE USE DIGITAL DATA TO MEASURE RISK?

Allen, G., & Harding, M. (2021). *Terrorism in Great Britain: The Statistics. https://bit.ly/3R90G57*

Commission for Countering Extremism. (2020). *How hateful extremists are exploiting the pandemic. https://bit.ly/3bTJiRE*

Conway, M., Scrivens, R., & Macnair, L. (2019). Right-Wing Extremists' Persistent Online Presence: History and Contemporary Trends. *ICCT Policy Brief. https://bit.ly/3AyDmYq*

Grieve, L. (2021). *Covid-19 Conspiracy in Ireland and the Far-Right Nexus. https://bit.ly/3Pa2BVd*

Member States Concerned By The Growing And Increasingly Transnational Threat of Extreme Right-Wing Terrorism. (2020). *CTED Trends Alert. https://bit.ly/3uoqwIo*

Robert Bowers Pittsburgh Synangogue Shootings. (2018). *NYTimes. https://nyti.ms/3OINbro*

Spence, S. (2020). Right-wing extremism: The new wave of global terrorism. *The Conversation. https://bit.ly/3akrWNo*

Williamson, B. Brenton Tarrant: the processes which brought him to engage in political violence. *CSTPV Short Papers. https://bit.ly/3nKJDJ3*

BURNS: A-Z OF TRUST

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior* 10, 123, 122-143. *https://bit.ly/3bNVhjz*

Burns, C., & Conchie, S. M. (2011). Measuring implicit trust and automatic attitude activation. In F. Lyon, G. Mollering, & M. Saunders (Eds.) *Handbook of Research Methods on Trust.* Edward Elgar, London.

Dietz, G., & Den Hartog, D. N. (2006). Measuring trust inside organisations. *Personnel Review,* 35, 557-588. *https://bit.ly/3af8rG1*

Earle, T. C., & Siegrist, M. (2008). Trust, Confidence and Cooperation Model: A framework for understanding the relation between trust and Risk Perception. *International Journal of Global Environmental Issues,* 8, 17-29. *https://bit.ly/3AvfAwA*

Ferrin, D. L., & Gillespie, N. A. (2009, June). Cultural differences and universals in the development of trust. In *22nd Annual IACM Conference Paper. https://bit.ly/3IjoVGA*

Lewicki, R. J., McAllister, D. J., & Bies, R. J. (1998). Trust and Distrust: New relationships and realities. *Academy of Management Review*, 23, 438-458. *https://bit.ly/3RiBSYu*

McAllister, D. J. (1995). Affect- and Cognition- based trust as foundations for interpersonal cooperation in organizations. *Academy of Management Journal*, 38, 24-59. *https://bit.ly/3OGPrPz*

Meyerson, D., Weick, K. E., & Kramer, R. M. (1996). Swift trust and temporary groups. In R. M. Kramer & T. R. Tyler (Eds.) *Trust in Organizations: Frontiers of Theory and Researc*h. Sage, Thousand Oaks.

Schoorman, D. F., Mayer, R. C. & Davies, J. H. (2007). An integrative model of organizational trust: Past, present and future. *Academy of Management Review,* 22, 344-354. *https://bit.ly/3NJbluJ*

Slovic, P. (1993). Perceived risk, trust and democracy. *Risk Analysis,* 13, 675-682. *https://bit.ly/3useoYe*

Zand, D. E. (1972). Trust and Mangerial Problem Solving. *Administrative Science Quarterly*, 17, 229-239. *https://bit.ly/3RbA2bJ*

CONCHIE & TAYLOR: TRUST SIGNALS

Adolphs, R. (2009). The social brain: Neural basis of social knowledge. *Annual Review of Psychology,* 60, 693-716. *https://bit.ly/3IdlaFH*

Bargh, J. A., & Ferguson, M. J. (2000). Beyond behaviorism: On the automaticity of higher mental processes. *Psychological Bulletin,* 126, 925-945. *https://bit.ly/3ApoRDi*

Bayliss, A. P., & Tipper, S. P. (2006). Predictive eye gaze cues and personality judgments: Should eye trust you? *Psychological Science*, 17, 514-520. *https://bit.ly/3IhANMr*

Donohue, W. A., & Druckman, D. (2008). Message framing surrounding the Oslo I Accords. *Journal of Conflict Resolution,* 53, 119-145. *https://bit.ly/3ONvF55*

Leander, N. P., Chartrand, T. L., & Bargh, J. A. (2012). You give me the chills: Embodies reactions to inappropriate amounts of behavioral mimicry. *Psychological Science,* 23, 772-779. *https://bit.ly/3bVeXSR*

Maatter, M., Truong, K. P., & Heylen, D. (2010) How turn-taking strategies influence users' impressions of an agent. In IVA (Ed.), *International conference on intelligent virtual agents*. Philadelphia, PA, USA.

Nicholson, S. (2017). *The relevant and reliable language theory: developing a language measure of trust for online groups.* Lancaster University. *https://bit.ly/3ylVhyP*

Olivola, C. Y., & Todorov, A. (2010). Elected in 100 milliseconds: Appearance based inferences and voting. *Journal of Nonverbal Behavior*, 34, 83-110. *https://bit.ly/3uqH2aS*

Pentland, A. (2008). *Honest signals: How they shape our world.* The MIT Press.

Porter, S. et al. (2010). Dangerous decisions: The impact of first impressions of trustworthiness. *Psychology, Crime and Law,* 16, 477-491. *https://bit.ly/3agjZIV*

GLIKSON: EMOTIONAL OVER-TRUST IN AI TECHNOLOGY

Ben Mimoun, M. S., Poncin, I., & Garnier, M. (2012). Case study—Embodied virtual agents: An analysis on reasons for failure. *Journal of Retailing and Consumer Services,* 19(6), 605–612. *https://bit.ly/3bSaslu*

Diederich, S., Brendel, A. B., & Kolbe, L. M. (2020). Designing Anthropomorphic Enterprise Conversational Agents. *Business and Information Systems Engineering,* 62(3), 193–209. *https://doi.org/10.1007/s12599-020-00639-y*

Durán, J. M., & Jongsma, K. R. (2021). Who is afraid of black box algorithm. *Journal of Medical Ethics,* 47, 329–335. *https://bit.ly/3RdJ1ZW*

Glikson, E., & Woolley, A. W. (2018). A Human-Centered Perspective on Human–AI Interaction: Introduction of the Embodiment Continuum Framework. *Collective Intelligence.*

Lambrecht, A., & Tucker, C. (2019). Algorithmic Bias? An Empirical Study of Apparent Gender-Based Discrimination in the Display of STEM Career Ads. *Management Science*, mnsc.2018.3093. *https://bit.ly/3uv7G2u*

Malle, B. F., Scheutz, M., Forlizzi, J., & Voiklis, J. (2016). Which robot am I thinking about? The impact of action and appearance on people's evaluations of a moral robot. *ACM/IEEE International Conference on Human-Robot Interaction,* 2016-April, 125–132. *https://doi.org/10.1109/HRI.2016.7451743*

Mirnig, N., Stollnberger, G., Miksch, M., Stadler, S., Giuliani, M., & Tscheligi, M. (2017). To err is robot: How humans assess and act toward an erroneous social robot. *Frontiers Robotics AI,* 4(MAY). *https://bit.ly/3yLdHKX*

Salem, M., Lakatos, G., Amirabdollahian, F., & Dautenhahn, K. (2015). Would you trust a (faulty) robot? *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction* - HRI '15, 141–148. *https://bit.ly/3ONujHs*

Waytz, A., Heafner, J., & Epley, N. (2014). The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle. *Journal of Experimental Social Psychology*, 52, 113–117. *https://doi.org/10.1016/j.jesp.2014.01.005*

HILLNER: RAPPORT AND TRUST: WHAT'S THE DIFFERENCE?

Gabbert, F., Hope, L., Luther, K., Wright, G., Ng, M., & Oxburgh, G. (2020). Exploring the use of rapport in professional information gathering contexts by systematically mapping the evidence base. *Applied Cognitive Psychology*, 35(2), 329-341. *https://doi.org/10.1002/acp.3762*

Macintosh, G. (2009). The role of rapport in professional services: antecedents and outcomes. *Journal of Services Marketing*, 23(2), 70-78. *https://bit.ly/3Ig1xge*

Neequaye, D. A., & Mac Giolla, E. (2021). The Use of the Term Rapport in the Investigative Interviewing Literature: A Critical Examination of Definitions. *PsyArXiV. https://doi.org/10.31234/osf.io/fmp8h*

Oleszkiewicz, J. S., & Granhag, P. A. (2019). Semi-cooperative sources' affective resistance and cognitive strategies. In R. Bull, & I. Blandón-Gitlin (Eds.), *The Routledge International Handbook of Legal and Investigative Psychology* (pp. 255-267). Routledge.

Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not So Different After All: A Cross-Discipline View Of Trust. *Academy of Management Review,* 23(3), 393-404. *https://bit.ly/3PaE5mW*

Tickle-Degnen, L., & Rosenthal, R. (1990). The Nature of Rapport and Its Nonverbal Correlates. *Psychological Inquiry,* 1(4), 285-293. *https://bit.ly/3acsR2s*

LEE: WHAT IS SIEGE CULTURE?

Jackson, P. (2020). Transnational Neo-Nazism in the USA, United Kingdom, and Australia. *Program on Extremism:* GW University. *https://bit.ly/3P4geVZ*

Johnson, B., & Feldman, M. (2021). Siege Culture After Siege: Anatomy of a neo-Nazi Terrorist Doctrine. *ICCT https://bit.ly/3OM41OZ*

Lee, B. & Knott, K. (2021). Fascist Aspirants: Fascist Forge and Ideological Learning in the Extreme-Right Online Milieu. *Behavioral Sciences of Terrorism & Political Aggression https://bit.ly/3OM5dsl*

Macklin, G. (2019). The Evolution of Extreme-Right Terrorism and Efforts to Counter It in the United Kingdom. *CTC Sentinel* 12(1) *https://bit.ly/3OOLNTE*

Ware, J. (2019). Siege: The Atomwaffen Division and Rising Far-Right Terrorism in the United States. *ICCT. https://bit.ly/3nDSVGN*

LESLIE & WELLS: EVALUATING TRUST & RAPPORT: A PRACTITIONER'S GUIDE

Abbe, A., & Brandon, S. E. (2014). Building and maintaining rapport in investigative interviews. Police Practice & Research: *An International Journal,* 15(3), 207–220. *https://bit.ly/3NrJGUk*

Alison, E. & Alison, L. (2020). *Rapport: The Four Ways to Read People.* Ebury.

Alison, L., Humann, M. & Waring, S. (2016). Building good rapport in interviews. *CREST Security Review*, 2, 6-7. *https://bit.ly/3AqlY8h*

Hanway, P. (2019). Cognitive Load at Interview: The Interviewer's Perspective. *CREST Security Review,* 9, 18-19. *https://bit.ly/3yLO7FJ*

Hillner, L. (2022). Rapport and trust: What's the difference? *CREST Security Review*, this issue.

Taylor, P. (2002). A Cylindrical Model of Communication Behaviour in Crisis Negotiations. *Human Communication Research*, 28 (1), 7-48. *https://bit.ly/3xZCdWZ*

Wells, S. (2018). Transitions in Negotiation: From Crisis to Success. *CREST Security Review,* 7, 6-9. *https://bit.ly/3ykC6FN*

CREST's Eliciting Information Framework has been designed to help practitioners navigate and use the existing behavioural and social science evidence base on eliciting information. You can read more about the Framework here: *https://bit.ly/3nFpBzw*

LOCKEY: RECOVERING FROM FAILURE: WHAT CAN SECURITY SERVICES DO TO REPAIR TRUST?

Bachmann, R., Gillespie, N., & Priem, R. (2015). Repairing trust in organizations and institutions: Toward a conceptual framework. *Organization Studies*, 36(9), 1123-1142. *https://doi.org/10.1177/0170840615599334*

Bradford, B., Yesberg, J. A., Jackson, J., & Dawson, P. (2020). Live facial recognition: Trust and legitimacy as predictors of public support for police use of new technology. *The British Journal of Criminology,* 60(6), 1502-1522. *https://doi.org/10.1093/bjc/azaa032*

Dirks, K. T., & de Jong, B. (2022). Trust within the workplace: A review of two waves of research and a glimpse of the third. *Annual Review of Organizational Psychology and Organizational Behavior,* 9(1), 247-276. *https://bit.ly/3Rdnq3P*

Eberl, P., Geiger, D., & Aßländer, M. S. (2015). Repairing trust in an organization after integrity violations: The ambivalence of organizational rule adjustments. *Organization Studies,* 36(9), 1205-1235. *https://bit.ly/3AqgYR3*

Forsyth, B. (2015). Banning bulk: Passage of the USA FREEDOM act and ending bulk collection. *Washington & Lee Law Review*, 72, 1307-1342.

Gillespie, N., Dietz, G., & Lockey, S. (2014). Organizational Reintegration and Trust Repair after an Integrity Violation: A Case Study. *Business Ethics Quarterly*, 24(3), 371-410. *https://doi.org/10.5840/beq2014437*

Gillespie, N., Lockey, S., Hornsey, M., & Okimoto, T. (2021). Trust Repair: A multilevel framework. In N. Gillespie, A. Fulmer, & R. J. Lewicki (Eds). *SIOP Organizational Frontiers Series.* New York: Routledge.

Jackson, J., Bradford, B., Stanko, B., & Hohl, K. (2012). *Just authority?: Trust in the police in England and Wales. London:* Willan.

Lewicki, R. J., & Brinsfield, C. (2017). Trust repair. *Annual Review of Organizational Psychology and Organizational Behavior,* 4(1), 287-313. *https://*

doi.org/10.1146/annurev-orgpsych-032516-113147

McEvily, B., Perrone, V., & Zaheer, A. (2003). Trust as an organizing principle. *Organization Science,* 14(1), 91-103. *https://bit.ly/3ynm8KS*

Mourtgos, S. M., Adams, I. T., & Nix, J. (2022). Elevated police turnover following the summer of George Floyd protests: A synthetic control study. *Criminology & Public Policy,* 21(1), 9-33. *https://bit.ly/3nK4wnM*

Pfarrer, M. D., Decelles, K. A., Smith, K. G., & Taylor, M. S. (2008). After the fall: Reintegrating the corrupt organization. *Academy of Management Review*, 33(3), 730-749. *https://doi.org/10.5465/amr.2008.32465757*

Verble, J. (2014). The NSA and Edward Snowden: Surveillance in the 21st century. *ACM Sigcas Computers and Society*, 44(3), 14-20. *https://doi.org/10.1145/2684097.2684101*

RADUCU: TRUST THY ENEMY: TRUST AND RELATIONSHIP-BUILDING BETWEEN SOURCE HANDLERS AND INFORMANTS

Billingsley, R. (2009). *Covert Human Intelligence Sources: The 'unlovely' Face of Police Work*. Waterside Press.

Brimbal, L., Kleinman, S. M., Oleszkiewicz, S., & Meissner, C. A. (2019). Developing rapport and trust in the interrogative context: An empirically-supported and ethical alternative to customary interrogation practices. In S.J. Barela, M.J. Fallon, G. Gaggioli, & J.D. Ohlin (Eds.), *Interrogation and torture: Integrating efficacy with law and morality* (pp. 141-196). Oxford University Press.

Dando, C. J., & Oxburgh, G. E. (2016). Empathy in the field: Towards a taxonomy of empathic communication in information gathering interviews with suspected sex offenders. *The European Journal of Psychology Applied to Legal Context,* 8(1), 27-33. *https://bit.ly/3NLsNUX*

Home Office (2021, January 11). *Guidance: Covert Human Intelligence Sources Bill Factsheet* (accessible version). Retrieved January 19, 2022, from Gov.UK: *https://bit.ly/3NN9fPO*

Jap, S. D., & Anderson, E. (2007). Testing a life-cycle theory of cooperative interorganizational relationships: Movement across stages and performance *Management Science,* 53(2), 260-275. *https://bit.ly/3OLnRRa*

Knapp, M. L. (1978). *Social intercourse: From greeting to goodbye*. Allyn & Bacon.

Knapp, M. L., Vangelisti, A. L., & Caughlin, J. P. (2014). *Interpersonal communication and human relationships* (7 ed.). Pearson.

Lewicki, R. J., Tomlinson, E. C., & Gillespie, N. (2006). Models of interpersonal trust development: Theoretical approaches, empirical evidence, and future directions. *Journal of Management*, 32(6), 991-1022. *https://bit.ly/3yUudbH*

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, 20(3), 709-734. *https://bit.ly/3RijYVM*

Moffett, L., Oxburgh, G. E., Dresser, P., Watson, S. J., & Gabbert, F. (2021). Inside the shadows: a survey of UK human source intelligence (HUMINT) practitioners, examining their considerations when handling a covert human intelligence source (CHIS). *Psychiatry, Psychology and Law*, 1-19. *https://bit.ly/3Avpa2z*

Nunan, J., Stanier, I., Milne, R., Shawyer, A., & Walsh, D. (2020). Eliciting human intelligence: police source handlers' perceptions and experiences of rapport during covert human intelligence sources (CHIS) interactions. *Psychiatry, Psychology and Law,* 27(4), 511-537. *https://bit.ly/3P3A83a*

Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so different after all: A cross-discipline view of trust. *Academy of Management Review,* 23(3), 393-404. *https://bit.ly/3bWFvmF*

Sunnafrank, M., & Ramirez Jr, A. (2004). At first sight: Persistent relational effects of get-acquainted conversations. *Journal of Social and Personal Relationships,* 21(3), 361-379. *https://bit.ly/3RdESVT*

SEARLE: TRUST = CONFIDENCE + VULNERABILITY. THE ROLE OF THE LEADER

Bandura, A. (2016). *Moral disengagement: How people do harm and live with themselves.* Worth Publishers, Macmillan Learning.

Braddick, O. J., Pickren, W. E., & Oxford University Press. (2016). *Oxford research encyclopedia of psychology*. *http://psychology.oxfordre.com/*

Colquitt, J. A., & Rodell, J. B. (2011). Justice, Trust, and Trustworthiness: A Longitudinal Analysis Integrating Three Theoretical Perspectives. *Academy of Management Journal*, 54(6), 1183–1206. *https://bit.ly/3R6WZwy*

Fida, R., Paciello, M., Tramontano, C., Fontaine, R. G., Barbaranelli, C., & Farnese, M. L. (2015). An Integrative Approach to Understanding Counterproductive Work Behavior: The Roles of Stressors, Negative Emotions, and Moral Disengagement. *Journal of Business Ethics,* 130(1), 131–144. *https://doi.org/10.1007/s10551-014-2209-5*

Jordan, J., Brown, M. E., Treviño, L. K., & Finkelstein, S. (2013). Someone to Look Up To: Executive–Follower Ethical Reasoning and Perceptions of Ethical Leadership. *Journal of Management,* 39(3), 660–683. *https://bit.ly/3urSRx1*

Knoll, M., Neves, P., Schyns, B., & Meyer, B. (2021). A Multi-Level Approach to Direct and Indirect Relationships between Organizational Voice Climate, Team Manager Openness, Implicit Voice Theories, and Silence. *Applied Psychology,* 70(2), 606–642. *https://doi.org/10.1111/apps.12242*

Kramer, R. M., & Pittinsky, T. L. (Eds.). (2012). *Restoring trust in organizations and leaders: Enduring challenges and emerging answers.* Oxford University Press.

Lapidot, Y., Kark, R., & Shamir, B. (2007). The impact of situational vulnerability on the development and erosion of followers' trust in their leader. *The Leadership Quarterly,* 18(1), 16–34. *https://bit.ly/3Av2Fuk*

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An Integrative Model of Organizational Trust. *The Academy of Management Review*, 20(3), 709. *https://doi.org/10.2307/258792*

Rice, C., & Searle, R. H. (2022). 'The Enabling Role of Internal Organizational Communication in Insider Threat Activity – Evidence From a High Security Organization'. *Management Communication Quarterly*, 08933189211062250. *https://doi.org/10.1177/08933189211062250*

Searle, R., Nienaber, A.-M. I., & Sitkin, S. B. (Eds.). (2017). *The Routledge companion to trust.* Routledge.

Searle, R.H. & C. Rice. (2018). Assessing and mitigating the impact of organisational change on counterproductive work behaviour: An operational (dis)trust based framework. Centre for Research and Evidence on Security Threats (CREST). *https://bit.ly/300sLyd*

Steinbach, A. L., Kautz, J., & Korsgaard, M. A. (2021). Caring for their own: How firm actions to protect essential workers and CEO benevolence influenced stakeholder sentiment during the COVID-19 pandemic. *Journal of Applied Psychology,* 106(6), 811–824. *https://doi.org/10.1037/apl0000928*

Weibel, A., Den Hartog, D. N., Gillespie, N., Searle, R., Six, F., & Skinner, D. (2016). How Do Controls Impact Employee Trust in the Employer? *Human Resource Management*, 55(3), 437–462. *https://bit.ly/3akN2ew*

Whitener, E. M., Brodt, S. E., Korsgaard, M. A., & Werner, J. M. (1998). Managers as Initiators of Trust: An Exchange Relationship Framework for Understanding Managerial Trustworthy Behavior. *The Academy of Management Review*, 23(3), 513. *https://doi.org/10.2307/259292*

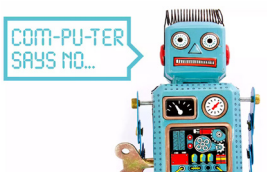
TAYLOR: TRUSTING A CENTRE MODEL

Edwards, J. (2019). Impact report. A review of the impact of the research projects conducted through the. Centre for Research and Evidence on Security Threats (CREST). *https://crestresearch.ac.uk/download/3427/impact-report-19-028-03.pdf*

PREVIOUS ARTICLES ON TRUST

Take a look at some of our past *CREST Security Review* articles that focus on, or feature, trust.

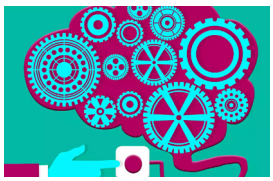
ASHENDEN: ALGORITHMIC DECISION MAKING (CSR#9)



How can redesigning system interactions help build trust between governments and citizens, enhance the security and wellbeing of individuals, and protect the security of the state?

https://crestresearch.ac.uk/comment/ashenden-algorithmic-decision-making/

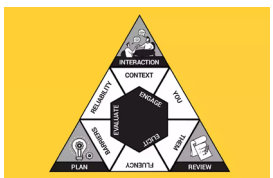
CIALDINI AND MARTIN: PERSUASION (CSR#8)



On the face of it, pre-suasion seems illogical. After all, how can we arrange for people to agree with a message before they know what's in it?

https://crestresearch.ac.uk/comment/power-of-persuasion-and-pre-suasion/

LESLIE: THE ELICITING INFORMATION FRAMEWORK (CSR#12)



Assisting practitioners in navigating the existing evidence base and more successfully applying it to their work.

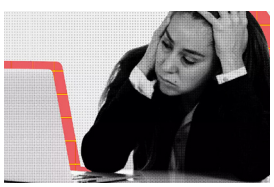
https://crestresearch.ac.uk/comment/the-eliciting-information-framework

MORRISON ET AL: THE ROLE OF (DIS)TRUST IN DISENGAGEMENT AND DERADICALISATION

When designing a disengagement or deradicalisation programme, who delivers it and how much they are trusted needs careful consideration.

https://crestresearch.ac.uk/comment/the-role-of-dis-trust-in-disengagement-and-deradicalisation/

NURSE: BALANCING CYBERSECURITY AND PRIVACY IN THE REMOTE WORKFORCE (CSR#12)



Remote working will only truly work if we get the balance of security and privacy right.

https://crestresearch.ac.uk/comment/balancing-cybersecurity-privacy-in-the-remote-workforce/

OLESZKIEWICZ: THE ADAPTABLE LAW ENFORCEMENT OFFICER (CSR#11)

What is adaptive behaviour? How can it be measured?

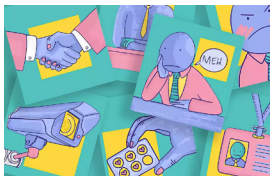
https://crestresearch.ac.uk/comment/the-adaptable-law-enforcement-officer/

OOSTINGA: COMMUNICATION ERROR HANDLING (CSR#6)

In Suspect Interviews And Crisis Negotiations we don't always make the correct decisions. How can we recover from different kinds of communication errors?

https://crestresearch.ac.uk/comment/communication-error-handling/

RICE AND SEARLE: TRUST IN ORGANISATIONAL CHANGE (CSR#8)



What are the strategies for mitigating the risk of insider threat from disillusioned employees?

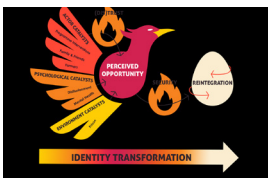
https://crestresearch.ac.uk/comment/positively-influencing-individuals-during-organisational-change/

SEARLE AND RICE: TRUST AND INSIDER THREAT (CSR#5)

How can networked trust in organisations, both between employees and processes, be maintained during times of great change?

https://crestresearch.ac.uk/comment/trust-insider-threat/

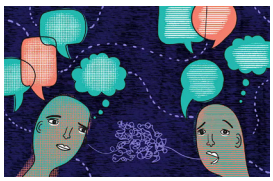
SILKE ET AL: THE PHOENIX MODEL (CSR#12)



Derived from a systematic review of contemporary research, a new model for understanding disengagement and deradicalisation processes has been produced.

https://crestresearch.ac.uk/comment/the-phoenix-model-disengagement-and-deradicalisation/

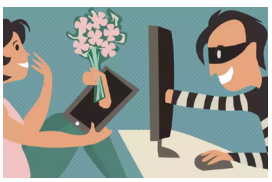
TAYLOR: COMMUNICATING ACROSS CULTURES (CSR#7)



From small talk to empathising, this article outlines some of the potential pitfalls and gaps in cross-cultural understanding.

https://crestresearch.ac.uk/comment/communicating-across-cultures/

WILLIAMS & JOINSON: TRUST AND ELICITING INFO ONLINE (CSR#1)



The internet often provides an ideal environment for those with malevolent intent to elicit information from victims.

https://crestresearch.ac.uk/comment/eliciting-information-online/

CENTRE FOR RESEARCH AND EVIDENCE ON SECURITY THREATS

CREST Security Review provides a gateway to the very best knowledge and expertise. Its articles translate academic jargon to 'so what' answers and illustrate how behavioural and social science can be used effectively in everyday scenarios.

THE CENTRE FOR RESEARCH AND EVIDENCE ON SECURITY THREATS

CSR is produced by the Centre for Research and Evidence on Security Threats (CREST). CREST is funded by the UK's Home Office and security and intelligence agencies to identify and produce social science that enhances their understanding of security threats and capacity to counter them. CREST also receives funding from its core partners (the universities of Bath, Lancaster and Portsmouth). Its funding is administered by the Economic and Social Research Council (ESRC Award ES/V002775/1), one of seven UK Research Councils, which direct taxpayers' money towards academic research and training. The ESRC ensures the academic independence and rigour of CREST's work.

CREST has established a growing international network of over 140 researchers, commissioned research in priority areas, and begun to tackle some of the field's most pressing questions.

> '*CREST Security Review is a fantastic means by which we can keep practitioners, policy-makers and other stakeholders up-to-date on the impressive social and behavioural science occurring not only at CREST, but around the world.*'
>
> Professor Stacey Conchie, CREST Director

For more information on CREST and its work visit **www.crestresearch.ac.uk** or find us on Twitter, @crest_research