

Human Factors

Sorting Insiders from Co-workers: Remote synchronous computer-mediated triage for investigating insider attacks

Journal:	<i>Human Factors : The Journal of the Human Factors and Ergonomics Society</i>
Manuscript ID	HF-21-7387.R2
Manuscript Type:	Research Article
Keywords:	Risk assessment < ACCIDENTS, HUMAN ERROR, Human-automation interaction < AUTOMATION, EXPERT SYSTEMS, Automatic and controlled processing < COGNITION, Dual task, time sharing, task switching < COGNITION, Language < COGNITION, Mental workload < COGNITION, Situated cognition < COGNITION

SCHOLARONE™
Manuscripts

Computer-mediated triage insider attacks

Sorting Insiders from Co-workers: Remote synchronous computer-mediated triage for investigating insider attacks

Coral J. Dando ¹, Paul J. Taylor ², Thomas C. Ormerod ³, Linden J. Ball ⁴, Alexandra L. Sandham ⁵, Tarek Menacere ²

¹ Dept. of Psychology, University of Westminster, London

² Dept. of Psychology, Lancaster University

³ School of Psychology, University of Sussex

⁴ School of Psychology, University of Central Lancashire

⁵ Dept. of Psychology, University of Gloucestershire

Acknowledgments: This research was funded by the UK Gov, Project Ref. LUR 46.851.

Correspondence should be addressed to:

Coral J. Dando,
Dept of Psychology,
University of Westminster, London.
Email: c.dando@westminster.ac.uk

Research Article

Key words: Insiders; Computer-mediated triage; Deception; Investigation

The data that support the findings of this study are available from the corresponding author upon reasonable request.

WORD COUNT: 4497 (excludes structured abstract, precis, bios, title page contents, key points, text in tables, figures and references)

Computer-mediated triage insider attacks

1

1 **Abstract**

2 **Objective:** Develop and investigate the potential of a remote, computer mediated and
3 synchronous text-based triage, which we refer to as *InSort*, for quickly highlighting persons
4 of interest after an insider attack.

5 **Background:** Insiders maliciously exploit legitimate access to impair the confidentiality and
6 integrity of organizations. The globalisation of organisations and advancement of information
7 technology means employees are often dispersed across national and international sites,
8 working around the clock, often remotely. Hence, investigating insider attacks is challenging.
9 However, the cognitive demands associated with masking insider activity offer opportunities.
10 Drawing on cognitive approaches to deception and understanding of deception-conveying
11 features in textual responses we developed InSort, a remote computer mediated triage.

12 **Method:** During a 6-hour immersive simulation, participants worked in teams, examining
13 password protected, security sensitive databases and exchanging information during an
14 organized crime investigation. Twenty-five percent were covertly incentivized to act as an
15 ‘insider’ by providing information to a provocateur.

16 **Results:** Responses to InSort questioning revealed insiders took longer to answer
17 investigation relevant questions, provided impoverished responses, and their answers were
18 less consistent with known evidence about their behaviors than co-workers.

19 **Conclusion:** Findings demonstrate InSort has potential to expedite information gathering and
20 investigative processes following an insider attack.

21 **Application:** InSort is appropriate for application by non-specialist investigators and can be
22 quickly altered as a function of both environment and event. InSort offers a clearly defined,
23 well specified, approach for use across insider incidents, and highlights the potential of
24 technology for supporting complex time critical investigations.

25

Computer-mediated triage insider attacks

2

26 **Precis**

27 Insiders exploit legitimate access to negatively affect organizations. Drawing on cognitive
28 approaches to deception and understanding of deception-conveying features in textual
29 responses we combined these literatures to develop 'InSort', a novel, remote computer-
30 mediated triage. Findings suggest InSort has potential to triage persons of interest from co-
31 workers following an attack, thereby expediting the initial investigative process.

32

Computer-mediated triage insider attacks

3

33 **Introduction**

34 Insiders exploit privileged access to damage organizations (see Mills et al., 2017; Posey et
35 al., 2013). Examples include a BUPA employee who downloaded and offered for sale
36 547,000 items of patient information and a NASA employee who downloaded classified
37 national defence information. Insider crime is increasing (Homoliak et al., 2019; Clearswift
38 Insider Threat Index, 2017) and becoming more expensive (European Union Agency for
39 Cybersecurity, 2020; National Law Review, 2020). Surveys suggest 27% of cybercrime
40 incidents are committed by insiders (Trzeciak, 2019) with insiders responsible for 43% of
41 data loss reported by the world's largest companies (Intel Security, 2015). Insider threats are
42 difficult to mitigate. Employees are trusted, with detailed knowledge and access to employer
43 assets. Understanding of insider behaviours and psychological characteristics is improving
44 (e.g., Costa, et al., 2016; Elmrabbit et al., 2020; Greitzer et al., 2018; Spitzner, 2003; Taylor et
45 al., 2013). However, few insider investigative techniques exist (Maybury, 2006) because
46 knowledge derived from one attack is not necessarily relevant to others (e.g., CPNI, 2020;
47 Saxena et al., 2020).

48 **Computer-Mediated Triage**

49 Gathering post attack information is fundamental to understanding what has
50 happened. In doing so, investigators (in-house security or external agencies) seek to
51 understand the veracity of employee accounts. Employees may be dispersed across numerous
52 national or international sites and so conducting timely and effective investigations can be
53 challenging. Here, we evaluate text-based computer mediated communication (CMC) using a
54 series of event-specific questions towards meeting this challenge. CMC screening is
55 increasingly used to support decision-making where there are high volumes of traffic such as
56 for pre-screening job applicants and completing employee credibility assessments (Jenson et
57 al., 2010; Tyman et al., 2014). Building on research concerning the language of insiders

Computer-mediated triage insider attacks

4

58 (Jenkins & Dando, 2011; Taylor et al., 2013), we investigated whether synchronous textual
59 responses to CMC questions might effectively triage persons of interest.

60 CMC has several potential advantages. Organizations can gather information from
61 employees simultaneously, irrespective of location, offering speed, volume, and reach (e.g.,
62 Lew et al., 2018; Pang et al., 2018; Yao & Ling, 2020). Text-based CMC is widely
63 accessible, technically stable and is low in media richness and so devoid of non-verbal cues
64 that occur during face-to-face interactions that can negatively impact investigations,
65 potentially reducing false positives and negatives (e.g., Bond & DePaulo, 2006; Dando &
66 Ormerod, 2017; Is baster & Nass, 2000; Markowitz , 2020; Matsumoto et al., 2011; Meissner
67 & Lyles, 2019; Nortje & Tredoux, 2019; ; Walsh et al., 2018).

68 Masking Malicious Behaviour

69 Psychological knowledge of the challenges of masking malicious activity offers
70 strategic insight into how to structure a CMC triage. To remain above suspicion necessitates
71 deceiving colleagues (e.g., Homoliak et al, 2019; Lew et al., 2018; Taylor et al., 2013).
72 Hence, insiders have an impression management goal (Colwell et al., 2006; Weiss et al.,
73 2006). They have to provide deceptive accounts that appear truthful and so have to manage
74 ‘two employment worlds’: tasks they should and should not have completed. Hence,
75 providing a convincing false account is more demanding than completing legitimate activity
76 and then providing a truthful account. This disparity offers opportunities for detection (e.g.,
77 Colwell et al., 2007; Kohan et al., 2020; Vrij et al., 2017).

78 Increased cognitive load in such circumstances (e.g., Bhatt et al., 2009; Jiang et al.,
79 2015) can result in differential verbal behaviours between liars and truth-tellers. Liars often
80 provide less consistent or coherent verbal accounts lacking informational content, with fewer
81 event details (Boggard et al., 2016; DePaulo et al., 2003; Hartwig et al. 2011). Differences
82 can be enhanced by tactical questioning techniques (e.g., Blandon-Gitlin et al., 2014; Dando

Computer-mediated triage insider attacks

5

83 & Bull, 2011; Dando & Ormerod, 2020; Hamlin et al., 2020; Ormerod & Dando, 2015;
84 Sporer, 2016; Vrij et al., 2010), which have yielded over 70% accuracy where the base rate of
85 deceivers was just 1:1000 (Dando & Ormerod, 2020; Ormerod & Dando, 2015), compared
86 with a typical detection rate of 54% (e.g., Bond & dePaulo, 2006; Hauch et al., 2016). Similar
87 results are reported in laboratory-based research (e.g., Dando & Bull, 2011; Granhag &
88 Hartwig, 2015; Levine, 2014; Sandham et al., 2020).

89 Detecting deception via tactical questioning is largely situated in face-to-face and
90 media-rich interview contexts. Nonetheless, several techniques lend themselves to CMC
91 triage with potential for leveraging measurable indicators of deception (Lee et al., 2009; Zhou
92 et al., 2002), particularly where comparisons can be made across employee responses
93 gathered following each insider attack (Burgoon et al., 2003; Rubin et al., 2015). For
94 example, deception-conveying features can sometimes include wordy replies with low
95 information (e.g., Pollina et al., 2017; Vendemia et al., 2005) and more expressions of
96 uncertainty (Zhou et al., 2002).

97 Towards a Solution

98 Combining cognitive approaches to deception and understanding of deception-
99 conveying features in textual responses, we developed a novel CMC text-based triage: *InSort*
100 (**Insider Sort**). InSort comprised a series of bespoke questions dictated by the insider event
101 itself, the run-up to the event, and workers day-to-day work activities (e.g., necessary,
102 unnecessary, and not allowed). Additionally, various questioning strategies were employed.
103 Target questions concern attack-specific behaviours, including behaviours in the run up to an
104 attack, questions about attempted access to databases, physical movements, and
105 communication. Target questioning increases cognitive complexity for insiders to maximize
106 the collection of triage-relevant information. Open questions (tell, explain, describe) gather
107 accounts about specific times, necessitating provision of expansive answers. These question

Computer-mediated triage insider attacks

6

108 types and their tactical presentation makes it challenging for insiders to provide a coherent
109 account (e.g., Dando & Ormerod, 2020; Dando & Bull 2011; Ormerod & Dando, 2015).

110 Target questions are manipulated to impose high cognitive demands on liars. They are
111 not presented *en bloc* nor chronologically, thereby introducing a temporal element (requiring
112 maintenance of six worlds – true and false versions of past, present and future). Some target
113 questions are repeated, accentuating between-question inconsistencies and contradictions,
114 which can be indicative of deceit (Blair et al., 2018; Chan & Bull, 2014; Vredeveldt et al.
115 2014). Responses are required before moving to the next question. Thus, InSort is interactive
116 (e.g., Lee et al., 2009; Sánchez-Junquera et al., 2020; Zhou et al., 2003), demanding higher
117 levels of cognitive engagement (Burgoon et al., 2010). The immediacy of InSort reduces
118 opportunities to construct deceptive accounts or confer with accomplices versus lengthier
119 triage processes conducted by human investigators (Levine & Blair, 2018; Walczyk et al.,
120 2013).

121 In sum, InSort may confer advantages including speed of implementation and
122 increased concurrent cognitive demand for insiders (deceivers), which may leverage
123 deception-conveying features (e.g., Bhatt et al., 2009; Jiang et al., 2015). We conducted a
124 ‘serious gaming’ empirical study, whereby participants were immersed in a full-day office-
125 based collaborative investigations of organized crime. The game, known as Confidential
126 Operations Simulation (iCOS: see Taylor et al., 2013), was played over a series of
127 competitive rounds. To establish a behavioural baseline, the first round was played with no
128 insider. In subsequent rounds, team members were assigned the role of ‘insider’, receiving
129 financial incentives to undertake illicit activities and not to be caught (see Method). The
130 study tested a series of hypotheses:

- 131 • Insiders will take significantly longer than non-insiders to complete InSort (H¹) because
132 of the dual impacts of tactical questioning and limited time to develop lie scripts.

Computer-mediated triage insider attacks

7

- 133 • Impression management will result in insider's text responses to open target questions
134 being shorter and with less information than non-insiders (H^2).
- 135 • Insiders will be less consistent in their responses to closed target questions, making
136 answer-evidence errors (H^3).
- 137 • Insiders will report finding InSort cognitively demanding and will be less confident in
138 their responses (H^4).

139

140

Method

141

Participants and Procedure

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

Sixty participants were paid £50 to take part in iCOS games lasting between 6 and 9

hours ($M = 6.8$ hours) - 26 males ($M_{age} = 25.67$, range 18 to 40 years), and 34 females ($M_{age} =$

23.8 years, range 19 to 30 years). Each game was split into four rounds and comprised 12

players, randomly assigned to a team (i) Fraud; (ii) Human Trafficking; and (iii) Narcotics.

Each team comprised four roles: Administrator, Field Agent, Intelligence Analyst and

Tactical Investigator. Status and responsibilities within teams were equal.

Teams had to solve a series of linked crimes, one in each round. Teams were

presented with intelligence updates about criminal gangs and used this information to guide

their searches of password-protected databases. Team players pieced together information to

identify gang members and their location. Players' database access was limited, so team

members worked together, exchanging information, recognizing connections across

databases, and engaging in collaborative problem solving. The team that most quickly

identified and located criminals 'won'. Teams were financially incentivized to win each

round (an additional £20 for winning the round). Each round lasted approximately ninety

minutes including regular breaks.

At the start, participants were randomly assigned to a team role. They received

instructions about the tasks to be completed, training on using the investigative databases,

Computer-mediated triage insider attacks

8

161 and familiarized themselves with the databases. To simulate a secure environment, players
162 worked in ‘silent’ offices, making notes using desktop publishing and spreadsheet software
163 and exchanging information using email, SMS messaging and mobile phone conversations.
164 They had access to a printer in a separate room. Once familiar with the environment, one
165 team member received instructions about the first crime to be investigated. All further
166 interaction with participants was conducted via email with ‘Gold Command’ (a confederate).
167 Gold Command issued instructions for subsequent tasks. By embedding task instructions into
168 the simulation, we hoped to enhance participants’ immersion in the simulation (Druckman,
169 2005).

170 Prior to the second round (at the end of the first round), up to two players in each
171 team were covertly approached to provide information to a provocateur for an additional £20
172 reward. Specifically, to obtain information concerning an individual under investigation, and
173 to email this information to the provocateur. The approach occurred face-to-face, out of sight
174 of the other participants. The same participants were again covertly invited to complete
175 further acts in the third and fourth rounds for an additional £20 each time. They were
176 instructed to develop their own method for completing the insider task to avoid raising
177 suspicion of teammates. All participants approached agreed to the insider task. The multiple
178 teams and sequence of rounds provided insiders numerous opportunities to complete their
179 tasks. For example, they could develop friendships with members of other teams for
180 malicious information gathering or distribute their activity across multiple periods to make it
181 more difficult to spot patterns of activity. Similarly, breaks taken by co-workers afforded
182 opportunities for players to compromise security.

183 Investigative tasks increased in complexity throughout the game. Similarly, the
184 insider task increased in complexity. In round 2, insiders were instructed to retrieve
185 information from a database they had legitimate access to but which was irrelevant to their

Computer-mediated triage insider attacks

9

186 team's intelligence task. In round 3, to provide information from a database only legitimately
187 accessible by another team member. In round 4, to gather information from a database that
188 was only accessible by members of another team. Once the game was complete, players were
189 informed that there had been a security breach, and that their behaviour during the simulation
190 would be investigated. Each participant was then required to individually complete InSort. All
191 insiders completed each of the insider tasks set.

192 Materials

193 The iCOS software comprised five primary modules: a password-protected database
194 creation module, a player interface, a data/keystroke capture module, an investigator
195 interface, and a game configuration module. The software provided an 'electronic' footprint
196 of activities undertaken by each player, including searches of particular databases, use of
197 email, use of internet, and use of printer for each system user. Footprint data and
198 communication data were used to verify participants' answers to InSort questions. Players
199 were informed that because they were working in a security sensitive environment they were
200 being monitored at all times. This included digital video recording, keystroke data, and
201 monitoring mobile phone usage (text and voice).

202 InSort comprised 56 questions, of which 16 were repeated (example questions see
203 appendix A):

- 204 • Two questions collected information regarding team membership and role, answered
205 via a drop-down menu.
- 206 • One question asked participants to indicate which databases they had access to as a
207 function of their role and team, again via a drop-down menu.
- 208 • Three open target questions invited textual responses regarding incident-specific
209 duties, communications activity and movements around the office including access to
210 the printer room and printing activity.

Computer-mediated triage insider attacks

- 211 • Eight forced-choice yes/no questions concerned password security and adherence to
212 iCOS rules and regulations regarding data security.

213 The following yes/no questions were repeated twice, randomly throughout the InSort
214 interview:

- 215 • Four related to access to each of the four databases.
- 216 • Four concerned attempted (but unsuccessful) access to each of the four databases.
217 Four concerned mobile phone usage (1), SMS messaging (1), emailing documents (1),
218 and email behaviour (1).
- 219 • Four questions concerned visiting the meeting room, meeting other players, visiting
220 the printer room, and printer use.

221 Participants received instructions on completing InSort, after which they logged in
222 using a unique identifier. Participants could only move forwards through InSort and were
223 unable to skip questions. On completion, participants provided feedback regarding player
224 strategies, behaviours and perceptions of InSort via a hard copy questionnaire comprising 10
225 questions with Likert scale (ranging from 1 to 5) or yes/no responses.

226 This research complied with the American Psychological Association Code of Ethics
227 and was approved by the Lancaster University Institutional Review Board. Informed consent
228 was obtained from each participant (materials are available from the first author).

229

Results

230 **Duration (H¹).** Two-way ANOVAs revealed a significant main effect of group
231 (insider, non-insider), $F(1, 54) = 187.81$ $p < .001$, $\eta_p^2 = 0.88$. Insiders took twice as long to
232 complete InSort ($M = 696s$, $SD = 120.28$, 95% CI, 626.62; 765.52) than non-insiders ($M =$
233 $340s$, $SD = 79.37$, 95% CI, 316.15; 363.29). Main effects of team (Narcotics, Fraud,
234 Trafficking) and team role (Administrator, Field Agent, Intelligence Analyst, Tactical
235 Investigator) and all interactions were non-significant, as were the all $F_s < 0.35$, all $p_s > .097$.

Computer-mediated triage insider attacks

11

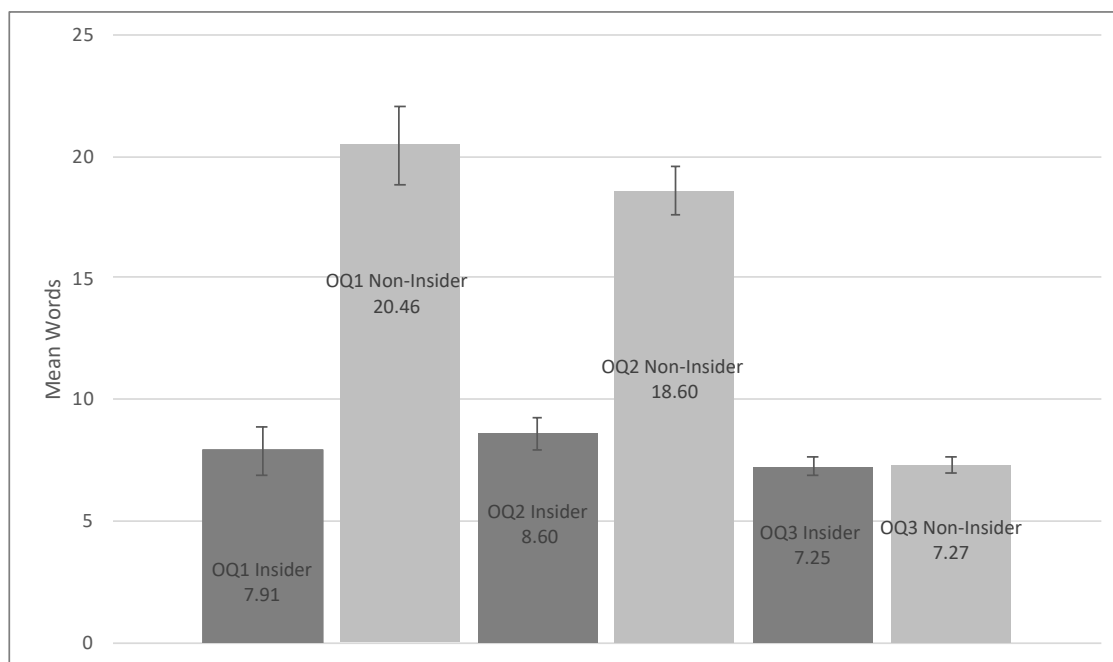
236 **Word count and information content (H²).** Two-way ANOVAs revealed a
237 significant main effect of group (insider, non-insider) for the total number of words in
238 response to each open target questions, $F(1, 36) = 12.866, p = .001, \eta_p^2 = 0.26$, and $F(1, 36)$
239 $= 23.95, p < .001, \eta_p^2 = 0.40$, respectively (see Fig. 1). Non-insiders wrote three times more
240 words (SD = 10.67) than insiders (SD = 2.21) for OQ1 and 2.5 more words (SD = 18.43) for
241 OQ2 than insiders (SD = 8.40). Main effects of team (Narcotics, Fraud, Trafficking) and
242 team role (Administrator, Field Agent, Intelligence Analyst, Tactical Investigator) were non-
243 significant, as were all interactions, all $ps > .554$ (see Table 1). OQ3 was only available to
244 participants who responded ‘yes’ to questions concerning printer usage, emailing documents
245 for printing and visiting the printer room. Accordingly, 25 participants responded to OQ3, of
246 which seven were insiders (50% of insiders; 30% of non-insiders). A one-way ANOVA
247 revealed no significant difference between insiders and non-insiders for total word count in
248 response to OQ3, $p = .894$ (see Fig. 1).

249

250 Figure 1.

251 *Mean word count for each of open question (OQ1, OQ2 and OQ3) as a function of group*
252 *(insider; non-insider).*

Computer-mediated triage insider attacks



253

254

255 Information items in response to open target questions (OQ1, OQ2 and OQ3) were
 256 calculated by summing the number of correct, discrete, quantifiable investigation relevant
 257 information (IRI) items (see Oxburgh et al., 2012; Philips et al., 2012 for more on IRI). For
 258 example, the following response was coded as six information items, *‘Over the day I was
 259 tasked with looking at conversations¹ and other intelligence information in the human
 260 trafficking intercepts database². I did this to try and track down and formulate an arrest list³
 261 for the leaders of the Zebra gang⁴, the Garfunkels gang⁵ and by working in collaboration
 262 with my team members, particularly the tactical investigator⁶’.*

263 Responses to open questions were initially coded by a researcher naïve to the research
 264 design and hypotheses following a set of guidelines. 20% (12) of responses from each of the
 265 three questions (randomly selected) then underwent independent secondary coding. Inter-
 266 rater agreement (IRA) between the coders was high for each of the open questions, $r = .916$
 267 (OQ1), $r = .882$ (OQ2) and $r = .902$ (OQ3).

268 Two-way ANOVAs revealed significant main effects of group for total information
 269 items in OQ1 (individual roles) and OQ2 (individual movements), $F(1, 36) = 9.485$, $p = .003$,

Computer-mediated triage insider attacks

13

270 $\eta_p^2 = 0.22$ and, $F(1, 36) = 34.75$, $p < .001$, $\eta_p^2 = 0.49$, respectively. No other main effects nor
 271 interactions emerged, all $ps > .071$. In response to OQ1 and OQ2, insiders provided far less
 272 information than non-insiders (see Table 1).

273

274 Table 1.

275 *Mean information items for each open question (OQ1, OQ2 and OQ3) as a function of group*
 276 *(insider; non-insider).*

	Insider	Non-Insider
	<i>M (95% CI)</i>	
Open Question 1	0.86 (.31: 1.41)	4.30 (3.43: 5.18)
Open Question 2	1.07 (.50: 1.65)	3.85 (3.43: 4.26)
Open Question 3	.86 (.22: 1.50)	1.50 (1.11: 1.89)

277

278 **Closed target question errors (H³).** Answers to each of the questions that comprised
 279 the four clusters of closed repeated target questions were scored as correct (awarded 1) or
 280 incorrect (awarded 2) at Time 1 (first presentation) and in a similar fashion again at Time 2
 281 (second presentation) resulting in an overall target question consistency score for each
 282 participant (lower score indicates fewer errors) per cluster (see Table 2). Answers were
 283 scored as correct only if participants responded in accordance with behaviours known to
 284 match the electronic footprint and surveillance data. The maximum error score (answered
 285 incorrectly at Time 1 & 2) was 16. A score of 8 indicated respondents were correct on both
 286 occasions.

287

288 Table 2.

289 *Mean target question cluster error scores a function of group (insider; non-insider) where,*
 290 *max. error score = 16, min. = 8.*

Computer-mediated triage insider attacks

14

291

292

293

294

295

296

297

	Insider	Non-Insider
	<i>M</i> (95% CI)	
Database Accessed	9.93 (9.07: 10.79)	9.45 (9.08: 9.88)
Database Access Attempted	10.43 (9.26: 11.60)	9.52 (9.09: 9.95)
Communication	12.36 (11.15: 13.57)	9.50 (9.05: 9.95)
Movement	8.71 (8.19: 9.24)	8.78 (8.05: 9.03)

298

299

300

301

302

303

304

305

306

Two-way ANOVAs revealed non-significant effects of group, team and team role and non-significant interactions for successful database access target questions, all $ps > .131$.

Similarly, target question scores for attempted database access revealed non significant main effects and interactions, all $ps > .077$. A significant main effect of group (insider, non-insider) emerged for target question scores for communication behaviours, $F(1, 36) = 29.268$, $p < .001$, $\eta_p^2 = 0.45$. Insider's scored higher than non-insiders', indicating discrepancies in responding. All other main effects and interactions were non-significant, all $ps > .103$. Target question scores for the cluster of movement questions revealed non-significant main effects and interactions, all $ps > .168$.

307

308

309

310

311

312

313

Answer-Evidence Inconsistency (H³). Answers to closed target questions at Time 1 were scored as consistent (1) or inconsistent (2) with known evidence. Scores were summed, referred to as the *answer-evidence inconsistency scale*, where a lower score indicates higher answer-evidence consistency. Mann-Whitney tests (data violated parametric assumptions) revealed a significant difference between insiders and non-insiders for answer-evidence inconsistency scores, $U = 43.00$, $z = -5.046$, $p < .001$, $r = -.084$. Overall, insiders' answer-evidence inconsistency scores were higher ($Mdn = 22.07$) than non-insiders' ($Mdn = 15.85$).

314

315

Post InSort Feedback (H⁴). All participants reported understanding the InSort instructions and complying with instructions. Insiders ($M_{Insiders} = 1.93$, 95% CI, 1.51: 2.35)

Computer-mediated triage insider attacks

15

316 reported InSort more demanding ($M_{\text{Non-Insiders}} = 3.52$, 95% CI, 3.27: 3.78), $F(1, 59) = 39.11$,
317 $p < .001$, $\eta_p^2 = 0.67$, and were less confident that their answers were correct, $F(1, 59) = 45.41$,
318 $p < .001$, $\eta_p^2 = 0.62$ ($M_{\text{Insiders}} = 4.29$, 95% CI, 3.52: 5.05 and $M_{\text{Non insiders}} = 2.04$, 95% CI,
319 1.74, 2.34). Insiders reported finding the questions more difficult, $F(1, 59) = 7.22$, $p = .009$,
320 $\eta_p^2 = 0.41$ ($M_{\text{Insiders}} = 1.50$, 95% CI, 1.20: 1.80 and $M_{\text{Non insiders}} = 2.11$, 95% CI, 1.87, 2.17).

321 Insiders reported being dishonest when answering questions, $X^2(1, 60) = 19.543$, $p <$
322 $.001$ and withholding information more often, $X^2(1, 60) = 24.65$, $p < .001$. There was no
323 difference between insiders and non-insiders when asked whether InSort questions had been
324 repeated, $p = .634$. Overall, 27 responded (45%) 'no', 16 (27%) were 'unsure' and 17 (28%)
325 said 'yes'. Again, no difference emerged between insiders and non-insiders as which types of
326 questions (yes/no or text responses) had been more demanding, $p = .370$. Overall, 25 (42%)
327 reported yes/no questions to be most demanding, 26 (43%) textual responses, and 9 (15%)
328 reported all questions were equally demanding.

Discussion

329 Insider attacks are increasing in number and magnitude, with potential to undermine
330 national and international security, cause financial loss and reputational damage (e.g., Legg,
331 2017; Wei et al., 2021). We developed InSort, a text-based synchronous triage with potential
332 for highlighting persons of interest after an insider incident. Insiders took twice as long to
333 complete InSort, were less confident their answers were correct, found InSort more
334 cognitively challenging, provided less information, and typed fewer words. Our results
335 confirm findings of previous research in face-to-face and remote person-to-person contexts
336 that questioning strategies which maximize cognitive burden can amplify signals of deception
337 (e.g., Boggard et al., 2016; DePaulo et al., 2003), highlighting the potential of remote
338 automated CMC.
339

Computer-mediated triage insider attacks

16

340 Open questions increased the information harvested, eliciting an individual's version
341 of the truth, which can be explored for veracity (e.g., Kontogianni et al., 2020; Snook et al.,
342 2010). Tactical questioning, concerning known or verifiable information are spread
343 throughout InSort rather than clustered at the beginning or end, which improves the veracity
344 performance by interviewers and observers (Dando et al., 2015; Levine, 2018). We
345 incorporated both where response time was not constrained, but where response time was
346 monitored. Yet, although respondents could take their time, did not have to consider social
347 context and how their answers/behaviours were received, tactical questioning again leveraged
348 diagnostic indicators across a cohort.

349 The remote CMC nature of InSort may have diverted impression management
350 towards behaviours perceived by insiders as more important, hence engendering differences
351 in the time taken to complete InSort and in the informational content in open question
352 responses. The absence of a human questioner, and without understanding the importance of
353 *all* response behaviours, some behaviours were attended to at the expense of others.
354 Providing a coherent and consistent narrative without contradictions, with little time to
355 prepare and where questions are not chronologically ordered, may explain the increased
356 duration. Insider responses to open target questions were shorter, suggesting they were
357 seeking to appear credible and cooperative, simultaneously being cautious in responding (see
358 Sporer, 2016; Schuetzler et al., 2019; Zukerman et al., 1981). Wordy replies with low
359 information can be indicative of deception, but not always. However, here short information
360 poor replies were indicative of insiders, possibly being deceptive by withholding information,
361 which is reported in face-to-face contexts (DeRosa et al., 2019; Levine, 2018)

362 Our findings are consistent with findings regarding the efficacy of automated
363 screening systems for detecting deception at border crossings and in job interviews, further
364 indicating that textual response content and response behaviours are important

Computer-mediated triage insider attacks

17

365 (Higginbotham, 2013; Nunamaker et al., 2011; Schuetzler, et al., 2019). Our results are also
366 consistent with cognitive load explanations of deceptive communication (Ho et al., 2016).
367 Creation and then typing of answers to questions is complex and time consuming, but the
368 additional demands associated with being deceptive is more time consuming still. Deceptive
369 textual communications are shorter due to the challenges of drawing multiple responses from
370 memory as plausible answers to questions (e.g., Burgoon et al., 2003; Pollina et al., 2017;
371 Schuetzler et al., 2019).

372 Manipulative questioning includes repeat questions, which we believed could
373 leverage notable inconsistencies between insiders and non-insiders because insiders would
374 struggle to provide credible and consistent responses to repeat questions (H³). Our question
375 cluster scores alone did not generally support this hypothesis. However, one important
376 finding was that insiders did not successfully monitor their communication behaviour and so
377 were unable to maintain consistency. Future triage approaches might consider capturing
378 detailed human-human remote interaction behaviours.

379 Although the consistency across time literature in face-to-face contexts is mixed, our
380 findings suggest deceivers can be as consistent, sometimes more so than truth-tellers (e.g.,
381 Blair et al., 2018; Clemens & Grolig, 2018; Masip et al., 2018). Conversely, answer-evidence
382 inconsistency scores differed significantly. While insiders were consistent in textual
383 responses, responses to target questions were inconsistent with evidence, which mirrors
384 results in face-to-face contexts (Hartwig et al., 2006; Sukumar et al., 2018). However, here
385 participants were aware their behaviour was monitored throughout and that movement
386 information was collected. In face-to-face contexts participants are often unaware of
387 information known by interviewers, which is fundamental to the success of tactical and
388 strategic interviewing techniques (e.g., see Oleszkiewicz & Watson, 2021). Here, despite

Computer-mediated triage insider attacks

18

389 knowing behaviour information was collected, answer-evidence inconsistency again emerges
390 as a useful metric with potential for improving veracity decisions.

391 Information Manipulation Theory 2 (McCornack et al., 2014: IMT2) may be relevant
392 whereby cognitive load is related to difficulty of reasoning through the problem space created
393 by a gap between the initial state, in our study the questions asked by InSort, and the end state
394 (avoidance of detection). IMT2 suggests lies are produced only when the production of the
395 truth is problematic, and that high cognitive load is not intrinsic to deceptive discourse but
396 depends on the potential number of solutions needed to present the version judged most
397 appropriate. Our game was designed to mimic demands experienced by insiders in a secure
398 environment. Hence, there were numerous narratives insiders could choose. IMT2 also
399 proposes quantity violations such as omitting problematic discourse as a frequent form of
400 deceptive discourse. This might explain why insiders produced fewer words.

401 Limitations and Future Directions

402 Our simulation embodied some features of organizations, but there are differences
403 between it and the real world. As Taylor et al. (2013) point out the absence of a ‘world’
404 outside the simulation as a limitation. Employees often communicate with individuals outside
405 their own organization, increasing the heterogeneity of communication and collaborative
406 behaviours. Insiders were chosen at random without controlling/measuring personality,
407 motivation, or personal circumstances, which may not tally with how insiders emerge. More
408 complex simulations could manage these variables. We compared known insiders to co-
409 workers as a first step towards understanding if InSort might leverage differences in textual
410 responses with reference to theories of cognitive load, information manipulation and
411 deception. More research is required to understand how to delineate signal from noise where
412 status is unknown. Finally, the structure of InSort is guided by the applied deception
413 literature and so likely to remain fairly consistent. However, the informational content of

Computer-mediated triage insider attacks

19

414 questions is dynamic. Ours was bespoke to the iCOS simulation. Constructing an event
415 specific InSort triage depends upon the nature of tasks workers are required and allowed to
416 do day-to-day, the information known to employers, and the insider event itself, which would
417 guide the informational content.

418 Conclusions

419 Findings demonstrate the potential of real time remote investigative triage approaches
420 such as InSort. InSort could regularly be implemented on an ad hoc basis as part of in-house
421 security practices following operations or investigations of the nature described here. This
422 may be useful for collating databases of response behaviours such as answer lengths and
423 response times. Such a database may offer additional information alongside the event specific
424 ‘footprint’ allowing comparisons across incidents. InSort can be constructed and
425 administered by non-specialists and quickly altered as required across incidents. As such,
426 InSort has potential to expedite investigative processes.

427

428 Key Points

- 429 • Investigating insider attacks is challenging because of the globalisation of
430 organisations and the fact that insiders exploit legitimate access.
- 431 • The acknowledged cognitive demands associated with masking illegal insider activity
432 offer opportunities.
- 433 • Drawing on cognitive approaches to deception and understanding of deception-
434 conveying features in textual responses we developed InSort, a rapid remote computer
435 mediated triage for highlighting persons of interest.
- 436 • InSort identified persons of interest and so could add to existing insider investigative
437 techniques following an insider attack.

Computer-mediated triage insider attacks

20

- 438 • InSort may be particularly relevant given the globalisation of organisations and
439 advancement of information technology whereby employees are dispersed across
440 national and international sites.

Computer-mediated triage insider attacks

21

441 **Appendix A: Example InSort Questions**

442 Example open question:

443 1. *'Please explain what your team role entailed'* Answer via free textual response

444 Example closed non-target questions:

445 1. *'What team were you assigned too?'* Answer via a forced choice (one choice allowed)

446 drop down menu

447 2. *'What was your role in the team?'* Answer via a forced choice (one choice allowed)

448 drop down menu

449 Example closed target questions (multiple responses option):

450 1. *'Which databases did your team role allow you to access?'* Answer via a drop down

451 menu allowing multiple choices

452 2. *'Which data bases did you access during the investigation?'* Answer via a drop down

453 menu allowing multiple choices

454 Example closed target questions (forced choice response):

455 1. *'Did you attempt to access the 'Shared Network' database?'* Yes/no456 2. *'Did you share your 'Shared Network' database password with anyone?'* Yes/no

457

458

459

460

461

462

463 **References**

- 464 Aamodt, M. G., & Custer, H. (2006). Who can best catch a liar? *Forensic Examiner*, 15(1), 6.
- 465 Bhatt, S., Mbwana, J., Adeyemo, A., Sawyer, A., Hailu, A., & Vanmeter, J. (2009). Lying
466 about facial recognition: an fMRI study. *Brain and Cognition*, 69(2), 382-390.
- 467 Blandón-Gitlin, I., Fenn, E., Masip, J., & Yoo, A. H. (2014). Cognitive-load approaches to
468 detect deception: searching for cognitive mechanisms. *Trends in Cognitive
469 Sciences*, 18(9), 441-444.
- 470 Blair, J. P., Reimer, T. O., & Levine, T. R. (2018). The role of consistency in detecting
471 deception: The superiority of correspondence over coherence. *Communication
472 Studies*, 69(5), 483-498.
- 473 Bogaard, G., Meijer, E. H., Vrij, A., & Merckelbach, H. (2016). Strong, but wrong: Lay
474 people's and police officers' beliefs about verbal and nonverbal cues to deception. *PloS
475 One*, 11(6), e0156615.
- 476 Bond, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and
477 Social Psychology Review*, 10(3), 214-234.
- 478 Burgoon, J. K., Floyd, K., & Guerrero, L. K. (2010). Nonverbal communication theories of
479 interpersonal adaptation. In *The New SAGE Handbook of Communication Science* (pp.
480 93-110). Sage.
- 481 Burgoon, J. K., Blair, J. P., Qin, T., & Nunamaker, J. F. (2003, June). Detecting deception
482 through linguistic analysis. In *International Conference on Intelligence and Security
483 Informatics* (pp. 91-101). Springer, Berlin, Heidelberg.
- 484 Chan, S., & Bull, R. (2014). The effect of co-offender planning on verbal
485 deception. *Psychiatry, Psychology and Law*, 21(3), 457-464.

Computer-mediated triage insider attacks

23

- 486 Chattopadhyay, P., Wang, L., & Tan, Y. P. (2018). Scenario-based insider threat detection
487 from cyber activities. *IEEE Transactions on Computational Social Systems*, 5(3), 660-
488 675.
- 489 Clearswift Insider Threat Index (2017) [https://www.clearswift.com/about-us/pr/press-](https://www.clearswift.com/about-us/pr/press-releases/insider-threat-74-security-incidents-come-extended-enterprise-not-hacking-groups)
490 [releases/insider-threat-74-security-incidents-come-extended-enterprise-not-hacking-](https://www.clearswift.com/about-us/pr/press-releases/insider-threat-74-security-incidents-come-extended-enterprise-not-hacking-groups)
491 [groups](https://www.clearswift.com/about-us/pr/press-releases/insider-threat-74-security-incidents-come-extended-enterprise-not-hacking-groups)
- 492 Clemens, F., & Grolig, T. (2019). Innocent of the crime under investigation: Suspects'
493 counter-interrogation strategies and statement-evidence inconsistency in strategic vs.
494 non-strategic interviews. *Psychology, Crime & Law*, 25(10), 945-962.
- 495 Colwell, L. H., Miller, H. A., Lyons Jr, P. M., & Miller, R. S. (2006). The training of law
496 enforcement officers in detecting deception: A survey of current practices and
497 suggestions for improving accuracy. *Police Quarterly*, 9(3), 275-290.
- 498 Colwell, K., Hiscock-Anisman, C. K., Memon, A., Taylor, L., & Prewett, J. (2007).
499 Assessment criteria indicative of deception (ACID): An integrated system of
500 investigative interviewing and detecting deception. *Journal of Investigative Psychology*
501 *and Offender Profiling*, 4(3), 167-180.
- 502 Costa, D. L., Albrethsen, M. J., & Collins, M. L. (2016). *Insider threat indicator ontology*.
503 <https://apps.dtic.mil/sti/citations/AD1044939>
- 504 CPNI (2020). *Investigation and Disciplinary*. [https://www.cpni.gov.uk/insider-](https://www.cpni.gov.uk/insider-risks/investigation-disciplinary)
505 [risks/investigation-disciplinary](https://www.cpni.gov.uk/insider-risks/investigation-disciplinary)
- 506 Dando, C. J., & Bull, R. (2011). Maximising opportunities to detect verbal deception:
507 training police officers to interview tactically. *Journal of Investigative Psychology and*
508 *Offender Profiling*, 8(2), 189-202.
- 509 Dando, C. J., & Ormerod, T. C. (2017). Analyzing decision logs to understand decision
510 making in serious crime investigations. *Human factors*, 59(8), 1188-1203.

Computer-mediated triage insider attacks

24

- 511 Dando, C. J., & Ormerod, T. C. (2020). Noncoercive human intelligence gathering. *Journal*
512 *of Experimental Psychology: General*, 149(8), 1435.
- 513 Dando, C. J., Bull, R., Ormerod, T. C., & Sandham, A. L. (2015). Helping to sort the liars
514 from the truth-tellers: The gradual revelation of information during investigative
515 interviews. *Legal and Criminological Psychology*, 20(1), 114-128.
- 516 DePaulo, B.M., Lindsay, J.J., Malone, B.E., Muhlenbruck, L., Charlton, K., & Cooper, H.
517 (2003). Cues to deception. *Psychological Bulletin*, 129(1), 74-118.
- 518 De Rosa, J., Hiscock-Anisman, C., Blythe, A., Bogaard, G., Hally, A., & Colwell, K. (2019).
519 A comparison of different investigative interviewing techniques in generating
520 differential recall enhancement and detecting deception. *Journal of Investigative*
521 *Psychology and Offender Profiling*, 16(1), 44-58.
- 522 Elmrabbit, N., Zhou, F., Li, F., & Zhou, H. (2020, June). Evaluation of machine learning
523 algorithms for anomaly detection. In *2020 International Conference on Cyber Security*
524 *and Protection of Digital Services (Cyber Security)* (pp. 1-8). IEEE.
- 525 European Union Agency for Cybersecurity (2020) Threat Landscape 2020 - Insider threat.
526 <https://www.enisa.europa.eu/publications/insider-threat>
- 527 Fenn, E., McGuire, M., Langben, S., & Blandón-Gitlin, I. (2015). A reverse order interview
528 does not aid deception detection regarding intentions. *Frontiers in Psychology*, 6, 1298.
- 529 Fisher, R. P., & Geiselman, R. E. (2010). The cognitive interview method of conducting
530 police interviews: Eliciting extensive information and promoting therapeutic
531 jurisprudence. *International Journal of Law and Psychiatry*, 33(5-6), 321-328.
- 532 Granhag, P. A., & Hartwig, M. (2015). The strategic use of evidence technique: A conceptual
533 overview. *Detecting deception: Current challenges and cognitive approaches*, 231-251.

Computer-mediated triage insider attacks

25

- 534 Greitzer, F., Purl, J., Leong, Y. M., & Becker, D. S. (2018, May). Sofit: Sociotechnical and
535 organizational factors for insider threat. In *2018 IEEE Security and Privacy Workshops*
536 (*SPW*) (pp. 197-206). IEEE.
- 537 Hamlin, I., Taylor, P. J., Cross, L., MacInnes, K., & Van der Zee, S. (2020). A psychometric
538 investigation into the structure of deception strategy use. *Journal of Police and*
539 *Criminal Psychology*, 1-11.
- 541 Hartwig, M., Granhag, P. A., Stromwall, L., Wolf, A. G., Vrij, A., & Hjelmsäter, E. R. A.
542 (2011). Detecting deception in suspects: Verbal cues as a function of interview
543 strategy. *Psychology, Crime & Law*, 17(7), 643-656.
- 544 Hartwig, M., Granhag, P. A., Strömwall, L. A., & Kronkvist, O. (2006). Strategic use of
545 evidence during police interviews: When training to detect deception works. *Law and*
546 *Human Behavior*, 30(5), 603-619.
- 547 Hauch, V., Sporer, S. L., Michael, S. W., & Meissner, C. A. (2016). Does training improve
548 the detection of deception? A meta-analysis. *Communication Research*, 43(3), 283-343.
- 549 Ho, S. M., Hancock, J. T., Booth, C., & Liu, X. (2016). Computer-mediated deception:
550 Strategies revealed by language-action cues in spontaneous communication. *Journal of*
551 *Management Information Systems*, 33(2), 393-420.
- 552 Homoliak, I., Toffalini, F., Guarnizo, J., Elovici, Y., & Ochoa, M. (2019). Insight into
553 insiders and it: A survey of insider threat taxonomies, analysis, modeling, and
554 countermeasures. *ACM Computing Surveys (CSUR)*, 52(2), 1-40.
- 555 Intel Security (2015). Grand Theft Data. Data exfiltration study: Actors, tactics, and
556 detection. https://cdn2.hubspot.net/hubfs/3375090/Nukon_June2017-theme/Pdf
557 [Files/rp-data-exfiltration.pdf](https://cdn2.hubspot.net/hubfs/3375090/Nukon_June2017-theme/Pdf)

Computer-mediated triage insider attacks

26

- 558 Jenkins, M. C., & Dando, C. J. (2012). Computer-mediated investigative interviews: A
559 potential screening tool for the detection of insider threat. In *Proceedings of the 10th*
560 *Biennial Conference of the International Conference of Forensic Linguistics,*
561 *Birmingham: Center for Forensic Linguistics.*
- 562 Jiang, W., Liu, H., Zeng, L., Liao, J., Shen, H., Luo, A., ... & Wang, W. (2015). Decoding the
563 processing of lying using functional connectivity MRI. *Behavioral and Brain*
564 *Functions, 11(1), 1-11.*
- 565 Jensen, C. S., Prasad, M. R., & Møller, A. (2013). Automated testing with targeted event
566 sequence generation. In *Proceedings of the 2013 International Symposium on Software*
567 *Testing and Analysis* (pp. 67-77).
- 568 Kohan, M. D., Nasrabadi, A. M., & Shamsollahi, M. B. (2020). Interview based connectivity
569 analysis of EEG in order to detect deception. *Medical Hypotheses, 136,* 109517.
- 570 Kontogianni, F., Hope, L., Taylor, P. J., Vrij, A., & Gabbert, F. (2020). “Tell me more about
571 this...”: An examination of the efficacy of follow-up open questions following an
572 initial account. *Applied Cognitive Psychology, 34(5), 972-983.*
- 573 Lee, C. C., Welker, R. B., & Odom, M. D. (2009). Features of computer-mediated, text-based
574 messages that support automatable, linguistics-based indicators for deception
575 detection. *Journal of Information Systems, 23(1), 5-24.*
- 576 Legg, P. A. (2017). Human-machine decision support systems for insider threat detection.
577 In *Data Analytics and Decision Support for Cybersecurity* (pp. 33-53). Springer, Cham.
- 578 Levine, T. R. (2018). Scientific evidence and cue theories in deception research: reconciling
579 findings from meta-analyses and primary experiments. *International Journal of*
580 *Communication, 12,* 19.
- 581 Levine, T. R. (2014). Truth-default theory (TDT) a theory of human deception and deception
582 detection. *Journal of Language and Social Psychology, 33(4), 378-392.*

Computer-mediated triage insider attacks

27

- 583 Levine, T. R., Blair, J. P., & Carpenter, C. J. (2018). A critical look at meta-analytic evidence
584 for the cognitive approach to lie detection: A re-examination of Vrij, Fisher, and Blank
585 (2017). *Legal and Criminological Psychology*, 23(1), 7-19.
- 586 Lew, Z., Walther, J. B., Pang, A., & Shin, W. (2018). Interactivity in online chat:
587 Conversational contingency and response latency in computer-mediated
588 communication. *Journal of Computer-Mediated Communication*, 23(4), 201-221.
- 589 Markowitz, D. M. (2020). The deception faucet: A metaphor to conceptualize deception and
590 its detection. *New Ideas in Psychology*, 59, 100816.
- 591 Masip, J., Martínez, C., Blandón-Gitlin, I., Sánchez, N., Herrero, C., & Ibabe, I. (2018).
592 Learning to detect deception from evasive answers and inconsistencies across repeated
593 interviews: A study with lay respondents and police officers. *Frontiers in*
594 *Psychology*, 8, 2207.
- 595 Maybury, M. (2006). *Detecting malicious insiders in military networks*. Mitre Corp. Bedford
596 MA. <https://apps.dtic.mil/sti/pdfs/ADA456254.pdf>
- 597 Matsumoto, D., Hwang, H. S., Skinner, L., & Frank, M. (2011). Evaluating truthfulness and
598 detecting deception. *FBI L. Enforcement Bull.*, 80, 1.
- 599 McCornack, S. A., Morrison, K., Paik, J. E., Wisner, A. M., & Zhu, X. (2014). Information
600 manipulation theory 2: A propositional theory of deceptive discourse
601 production. *Journal of Language and Social Psychology*, 33(4), 348-377.
- 602 Meissner, C. A., & Lyles, A. M. (2019). IX investigations: The importance of training
603 investigators in evidence-based approaches to interviewing. *Journal of Applied*
604 *Research in Memory and Cognition*, 8(4), 387-397.
- 605 Mills, J. U., Stuban, S. M., & Dever, J. (2017). Predict insider threats using human
606 behaviors. *IEEE Engineering Management Review*, 45(1), 39-48.

Computer-mediated triage insider attacks

28

- 607 National Law review (2020). Frequency and Cost of Insider Threats Continue to
608 Increase. [https://www.natlawreview.com/article/frequency-and-cost-insider-threats-](https://www.natlawreview.com/article/frequency-and-cost-insider-threats-continue-to-increase)
609 [continue-to-increase](https://www.natlawreview.com/article/frequency-and-cost-insider-threats-continue-to-increase)
- 610 Nortje, A., & Tredoux, C. (2019). How good are we at detecting deception? A review of
611 current techniques and theories. *South African Journal of Psychology*, 49(4), 491-504.
- 612 Ormerod, T. C., & Dando, C. J. (2015). Finding a needle in a haystack: Toward a
613 psychologically informed method for aviation security screening. *Journal of*
614 *Experimental Psychology: General*, 144(1), 76.
- 615 Oleszkiewicz, S., & Watson, S. J. (2021). A meta-analytic review of the timing for disclosing
616 evidence when interviewing suspects. *Applied Cognitive Psychology*, 35(2), 342-359
- 617 Oxburgh, G., Ost, J., & Cherryman, J. (2012). Police interviews with suspected child sex
618 offenders: does use of empathy and question type influence the amount of investigation
619 relevant information obtained?. *Psychology, Crime & Law*, 18(3), 259-273.
- 620 Pang, A., Shin, W., Lew, Z., & Walther, J. B. (2018). Building relationships through dialogic
621 communication: organizations, stakeholders, & computer-mediated
622 communication. *Journal of Marketing Communications*, 24(1), 68-82.
- 623 Parkhouse, T.P. and Ormerod, T.C. (2018). Unanticipated questions can yield unanticipated
624 outcomes in investigative interviews *PLoS ONE*, 13 (12). e0208751.
- 625 Pentland, S. J., Twyman, N. W., Burgoon, J. K., Nunamaker Jr, J. F., & Diller, C. B. (2017).
626 A video-based screening system for automated risk assessment using nuanced facial
627 features. *Journal of Management Information Systems*, 34(4), 970-993.
- 628 Phillips, E., Oxburgh, G., Gavin, A., & Myklebust, T. (2012). Investigative interviews with
629 victims of child sexual abuse: The relationship between question type and investigation
630 relevant information. *Journal of Police and Criminal Psychology*, 27(1), 45-54.

Computer-mediated triage insider attacks

29

- 631 Pollina, D. A., Woods, R. J., Salyer, C. D., Leffingwell, T. G., Cooper, C., & Rohrbaugh, J.
632 W. (2017). Verbal response time and duration indices of deception in humans
633 interviewed by a computer-generated agent. *International Journal of Human-Computer*
634 *Studies*, 97, 23-33.
- 635 Posey, C., Roberts, T. L., Lowry, P. B., Bennett, R. J., & Courtney, J. F. (2013). Insiders'
636 protection of organizational information assets: Development of a systematics-based
637 taxonomy and theory of diversity for protection-motivated behaviors. *Mis Quarterly*,
638 1189-1210.
- 639 Rubin, V. L., Conroy, N. J., & Chen, Y. (2015). Towards news verification: Deception
640 detection methods for news discourse. In *Hawaii International Conference on System*
641 *Sciences* (pp. 5-8).
- 642 Sánchez-Junquera, J., Villasenor-Pineda, L., Montes-y-Gómez, M., Rosso, P., & Stamatatos,
643 E. (2020). Masking domain-specific information for cross-domain deception
644 detection. *Pattern Recognition Letters*, 135, 122-130.
- 645 Sandham, A. L., Dando, C. J., Bull, R., & Ormerod, T. C. (2020). Improving professional
646 observers' veracity judgements by tactical interviewing. *Journal of Police and*
647 *Criminal Psychology*, 1-9.
- 648 Saxena, N., Hayes, E., Bertino, E., Ojo, P., Choo, K. K. R., & Burnap, P. (2020). Impact and
649 key challenges of insider threats on organizations and critical
650 businesses. *Electronics*, 9(9), 1460.
- 651 Schuetzler, R. M., Grimes, G. M., & Giboney, J. S. (2019). The effect of conversational agent
652 skill on user behavior during deception. *Computers in Human Behavior*, 97, 250-259.
- 653 Spitzner, L. (2003). Honeypots: Catching the insider threat. In *19th Annual Computer*
654 *Security Applications Conference, 2003. Proceedings.* (pp. 170-179). IEEE.

Computer-mediated triage insider attacks

30

- 655 Sporer, S. L. (2016). Deception and cognitive load: Expanding our horizon with a working
656 memory model. *Frontiers in Psychology*, 7, 420.
- 657 Sukumar, D., Wade, K. A., & Hodgson, J. S. (2018). Truth-tellers stand the test of time and
658 contradict evidence less than liars, even months after a crime. *Law and Human*
659 *Behavior*, 42(2), 145.
- 660 Taylor, P. J., Dando, C. J., Ormerod, T. C., Ball, L. J., Jenkins, M. C., Sandham, A., &
661 Menacere, T. (2013). Detecting insider threats through language change. *Law and*
662 *Human Behavior*, 37(4), 267.
- 663 Trzeciak, R. (2019). *SEI Cyber Minute: Insider Threat Mitigation, We can help!*. Carnegie
664 Mellon University Software Engineering Institute Pittsburgh United States.
- 665 Twyman, N. W., Lowry, P. B., Burgoon, J. K., & Nunamaker Jr, J. F. (2014). Autonomous
666 scientifically controlled screening systems for detecting information purposely
667 concealed by individuals. *Journal of Management Information Systems*, 31(3), 106-
668 137.
- 669 Vendemia, J., Buzan, R. F., & Simon-Dack, S. L. (2005). Reaction time of motor responses
670 in two-stimulus paradigms involving deception and congruity with varying levels of
671 difficulty. *Behavioural Neurology*, 16(1), 25-36.
- 672 Vredeveltdt, A., van Koppen, P. J., & Granhag, P. A. (2014). The inconsistent suspect: A
673 systematic review of different types of consistency in truth tellers and
674 liars. *Investigative Interviewing*, 183-207.
- 675 Vrij, A., Granhag, P. A., & Porter, S. (2010). Pitfalls and opportunities in nonverbal and
676 verbal lie detection. *Psychological Science in the Public Interest*, 11(3), 89-121.
- 677 Vrij, A., Meissner, C. A., Fisher, R. P., Kassin, S. M., Morgan III, C. A., & Kleinman, S. M.
678 (2017). Psychological perspectives on interrogation. *Perspectives on Psychological*
679 *Science*, 12(6), 927-955.

Computer-mediated triage insider attacks

31

- 680 Wei, Y., Chow, K. P. P., & Yiu, S. M. (2021). Insider Threat Prediction Based on
681 unsupervised Anomaly Detection Scheme for Proactive Forensic Investigation. *Digital*
682 *Investigation*.
- 683 Walczyk, J. J., Griffith, D. A., Yates, R., Visconte, S., & Simoneaux, B. (2013). Eye
684 movements and other cognitive cues to rehearsed and unrehearsed deception when
685 interrogated about a mock crime. *Applied Psychology in Criminal Justice*, 9(1), 1-23.
- 686 Walsh, D., Dando, C. J., & Ormerod, T. C. (2018). Triage decision-making by welfare fraud
687 investigators. *Journal of Applied Research in Memory and Cognition*, 7(1), 82-91.
- 688 Weiss, B., & Feldman, R. S. (2006). Looking good and lying to do it: Deception as an
689 impression management strategy in job interviews. *Journal of Applied Social*
690 *Psychology*, 36(4), 1070-1086.
- 691 Yao, M. Z., & Ling, R. (2020). "What is computer-mediated communication?" *Journal of*
692 *Computer-Mediated Communication*, 25(1), 4-8.
- 693 Zhou, L., Burgoon, J. K., Nunamaker, J. F., & Twitchell, D. (2004). Automating linguistics-
694 based cues for detecting deception in text-based asynchronous computer-mediated
695 communications. *Group Decision and Negotiation*, 13(1), 81-106.
- 696 Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal
697 communication of deception. *Advances in Experimental Social Psychology*, 14, pp. 1-
698 59.

Computer-mediated triage insider attacks

32

699 **Biographies:**

700

701 Prof Coral J. Dando is a Professor of Forensic Psychology, at the University of Westminster,
702 London. She has a PhD in Applied Cognition, awarded in 2011 by London South Bank
703 University. Prior to completing her PhD, Coral served as a London police officer for over 10
704 years.

705

706 Prof. Paul J. Taylor is a Professor of Psychology at Lancaster University, UK He has a PhD
707 in Psychology, awarded in 2004 by the University of Liverpool, UK. He is currently the
708 National Scientific Advisor for Policing. Paul was previously the Director of the UK's hub
709 for behavioural and social science for national security.

710

711 Prof Thomas C. Ormerod is a Professor of Psychology at the University of Sussex, UK.
712 He has a PhD in Human Computer Interaction, awarded in 1988 by the University of
713 Sunderland, UK. Tom has previously been Head of the School of Psychology at University of
714 Sussex.

715

716 Prof Linden Ball is a Professor of Psychology in the School of Psychology and Computer
717 Science at the University of Central Lancashire (UCLan). He has a PhD in Cognitive
718 Processes in Engineering Design, awarded in 1988 by the South West Polytechnic, Plymouth.
719 Linden is Director of Research and Enterprise in the Faculty of Science and Technology and
720 Deputy Director of the UCLan Research Centre for Brain and Behaviour.

721

722 Dr Alexandra Sandham is a Senior Lecturer in the Psychology Dept. at the University of
723 Gloucestershire, UK. She has a PhD in Hypothesis Generation in Investigative Contexts,

Computer-mediated triage insider attacks

33

724 awarded 2012 by Lancaster University. Alexandra has previously worked as a Principal

725 Psychologist at the UK Ministry of Defence.

726

727 Mr Tarek Menacere is a software developer. He is currently a Software Developer for Sky.

728 Tarek was previously employed at Lancaster University.