ThreatGEN® AutoTableTop 2.0

2026 WATER UTILITY COMPETITION

Pre-Competition Study

Abstract

ThreatGEN AutoTableTop™ 2.0 is an AI-powered tabletop exercise platform that enhances crisis response readiness through realistic cyber and operational scenarios. This report analyzes twenty ransomware exercise simulations, highlighting key response patterns, performance outcomes, and after-action insights. The findings demonstrate how structured decision-making, stakeholder coordination, and continuous improvement contribute to stronger organizational resilience and more effective incident response.

Robert C. Rhodes
robert@threatgen.com

Executive Summary

This document presents the integrated findings from a 20-trial study of AutoTableTop™ 2.0 running a Disaster Recovery crisis management tabletop exercise for a Texas water and wastewater treatment facility. The analysis combines two data sources: the raw facilitator transcripts (20 exercise sessions with per-turn scoring by the automated judge) and the corresponding After-Action Reports (AARs) generated by AutoTableTop™ 2.0 at the conclusion of each exercise.

The scenario simulates a ransomware attack discovered on the facility's Human-Machine Interface (HMI) at the start of a morning shift. Participants, primarily water plant operators, must detect, contain, and recover from the incident while maintaining water safety and public trust.

Key Findings

Across all 20 trials, zero casualties were recorded, no evacuations were required, and water supply was maintained through manual operations in every instance.

The optimal demo script consists of 5–6 well-crafted facilitator inputs. Trials achieving the highest average scores (7.0–7.5) shared a consistent five-phase decision framework.

The AARs revealed recurring improvement themes: public communication strategy, proactive engagement during resolution phases, and the critical importance of pre-established DFIR contracts. Only one trial (A12) triggered MITRE ATT&CK tactic identification (Initial Access, Execution, Impact), suggesting the AAR engine's TTP mapping is scenario-dependent and reserved for runs with sufficient technical depth.

Methodology

The study analyzed 20 complete exercise transcripts (designated A1 through A20), each representing a full run of the same ransomware scenario, plus 20 corresponding After-Action Reports automatically generated by AutoTableTop™ 2.0 upon exercise conclusion.

Transcript Analysis Dimensions

- **Facilitator Input Count**: Total participant submissions per trial (range: 3 to13)
- **Per-Turn Scores**: Automated judge scored each input on a 0 to 10 scale based on crisis management principles
- **Average Score**: Mean of all per-turn scores within a trial
- **Input Phrasing Analysis**: Specific wording patterns correlated with higher or lower scores
- **Narrative Efficiency**: Whether the trial reached clean resolution without stalling

After-Action Report Analysis Dimensions

- **Crisis Response Metrics**: Casualties, evacuations, time to critical decisions, resource utilization, public communication effectiveness, multi-agency coordination score, overall containment success
- **Successful Actions**: Highest-scoring actions identified by the AAR engine with turn references
- **Improvement Areas**: Lowest-scoring actions flagged for remediation
- **Operational Learnings**: Validated takeaways from the exercise
- **Safety Recommendations**: Forward-looking remediation steps
- **MITRE ATT&CK Tactic Mapping**: Adversary TTPs identified during the scenario

Scoring Summary by Trial

The table below summarizes performance across all 20 trials, integrating transcript scores with AAR-reported multi-agency coordination and overall containment assessments:

Trial	Inputs	Avg Score	High	Low	Coord. Score	Containment
A1	10	5.6	7	5	6/10	Fair
A2	12	6.2	8	0	Moderate	High
A3	7	7.0	8	6	High	High
A4	10	6.6	8	4	N/A	N/A
A5	5	7.2	8	6	7/10	High
A6	5	6.8	8	6	N/A	N/A
A7	6	6.3	8	5	N/A	N/A
A8	3	6.7	7	6	N/A	N/A
A9	8	6.0	8	0	N/A	N/A
A10	13	5.6	7	1	N/A	N/A
A11	6	7.2	8	7	7/10	7.2/10
A12	8	6.2	8	3	Moderate	Good
A13	6	6.8	7	6	N/A	N/A
A14	7	7.1	8	5	High	Successful
A15	4	7.0	8	6	N/A	N/A
A16	7	7.1	8	6	High	High
A17	6	5.7	8	2	N/A	N/A
A18	6	6.0	7	1	N/A	N/A
A19	4	7.5	8	7	High	Successful
A20	3	6.7	7	6	N/A	N/A

Consolidated After-Action Report Findings

Crisis Response Metrics (Aggregated Across 20 Trials)

The following metrics were consistent across all 20 AARs:

- **Total Casualties**: 0 across all 20 trials. Every trial prevented casualties through timely manual operations.
- **Evacuations**: None required in any trial. Manual operations maintained water safety throughout.

- **Time to Critical Decisions**: Ranged from immediate (within minutes) to 4 hours. Top-performing trials achieved critical decisions within 30 minutes of incident discovery.
- **Resource Utilization**: Consistently rated "*High*" or "*Efficient*" across trials, particularly in DFIR engagement and manual operations.
- **Public Communication Effectiveness**: The most variable metric across trials. Rated "*High*" in trials with proactive public statement preparation (A2, A3, A11, A16) but "*Moderate*" or "*Needs Improvement*" in trials that deferred communications (A5, A15, A19).
- **Multi-Agency Coordination Score**: Ranged from 6/10 (A1) to "*High*" (A3, A14, A16, A19). The score correlated strongly with early DFIR engagement and explicit authority notification.
- **Overall Crisis Containment**: Rated "*High*" or "*Successful*" in all trials that achieved an average transcript score of 7.0 or above.

Universally Successful Actions

The AARs identified the following actions as successful in every trial where they were attempted:

1. **Notifying authorities and switching to manual operations** (Turn 1 in all 20 trials, Score: 7 to 8). This was the single most consistent high-scoring action across the entire study. Every AAR cited it as a successful action.
2. **Engaging DFIR / cybersecurity incident response resources** (Turn 2 in most trials, Score: 7 to 8). The AARs from A2, A4, A6, A12, and A16 specifically highlighted the value of pre-existing DFIR contracts. A7's AAR noted the absence of such a contract as an improvement area.
3. **Preparing public statements and engaging legal stakeholders** (Score: 7 to 8 when combined in a single action). AARs from A3, A9, A14, and A15 scored this action 7 to 8 and flagged it as essential for maintaining public trust.
4. **Preparing lessons learned documentation and conducting debrief** (Score: 7 to 8). AARs from A2, A3, A5, A11, and A19 explicitly identified this as a successful action that supports continuous improvement.

Recurring Improvement Areas

The AARs flagged the following patterns as areas needing improvement:

- **Passive waiting during resolution phases:** The single most common improvement area across all AARs. Trials A1, A4, A9, A10, A12, A13, A17, and A18 were specifically flagged for passive responses ("OK," "await," "continue") during the cybersecurity team's remediation work. The AARs consistently recommended proactive engagement rather than passive monitoring.
- **Public communication and misinformation management:** AARs from A3, A5, A11, A15, A16, A17, A19, and A20 identified public communication as needing improvement. The scenario's adversary agent actively spreads misinformation through social media in every trial, and the AARs noted that participants rarely addressed this vector directly.
- **Transition planning (manual to automatic):** AARs from A6, A8, A12, and A20 flagged the transition back to automated operations as requiring more careful evaluation. The recommendation was to verify system integrity with the DFIR team before reintroducing the HMI, rather than assuming clearance.
- **Lack of detailed action plans:** AARs from A4, A7, and A17 noted that vague delegation ("assist the team," "follow through with next steps") produced lower scores. The recommendation was to specify concrete actions rather than abstract support.

Consolidated Operational Learnings

Synthesizing the operational learnings reported across all 20 AARs:

1. **Manual operations as contingency**: Multiple AARs (A5, A8, A10, A16, A19, A20) validated that pre-trained manual operation capability is essential for maintaining service during cyber incidents. This was the most frequently cited operational learning.
2. **DFIR contract readiness**: AARs from A2, A7, A12, A17, A18 emphasized that pre-established DFIR contracts dramatically reduce response time. A18's AAR explicitly recommended establishing a DFIR contract as its primary safety recommendation.
3. **Communication protocol enhancement**: AARs from A2, A8, A17, A18, A20 identified that rumor management and clear communication channels require dedicated protocols, not ad hoc responses.
4. **Proactive cybersecurity team engagement**: AARs from A2, A5, A10, A12 found that actively supporting the DFIR team (providing access, data, context) accelerates resolution compared to passive waiting.
5. **Incident logging for post-crisis analysis**: A20's AAR specifically recommended maintaining detailed incident logs throughout the crisis to support the post-incident debrief and future preparedness.

Consolidated Safety Recommendations

The following safety recommendations appeared across multiple AARs and represent the consensus findings:

1. **Regular software updates and patch management** (cited in A1, A10, A12, A17, A19, A20): Prevent exploitation of known vulnerabilities, particularly on legacy SCADA/HMI systems.
2. **Phishing awareness and prevention training** (cited in A2, A16): Multiple AARs identified phishing as the suspected attack vector. Regular simulation exercises were recommended.
3. **Establish or maintain DFIR contracts** (cited in A7, A18): Ensure rapid expert response capability exists before an incident occurs.
4. **Regular cybersecurity drills and tabletop exercises** (cited in A3, A8, A16, A18, A20): Practice manual operation transitions, incident response procedures, and cross-team coordination.
5. **Robust public communication strategies** (cited in A3, A5, A11, A15, A16, A17, A19): Develop pre-drafted communication templates, designate spokespersons, and establish social media monitoring for misinformation.
6. **Backup and recovery system testing** (cited in A1, A20): Regularly test backup systems and recovery procedures to ensure they function when needed.
7. **Advanced threat detection** (cited in A11, A19): Implement early-warning detection mechanisms to identify ransomware before it encrypts critical systems.

MITRE ATT&CK TTP Analysis

Of the 20 AARs, only Trial A12 triggered the MITRE ATT&CK tactic identification engine. The remaining 19 AARs reported "*None*" for MITRE Tactics. This pattern provides important insight into AutoTableTop™ 2.0's AAR generation behavior.

A12 Identified Tactics

Trial A12's AAR identified three MITRE ATT&CK tactics:

Tactic	MITRE ID	Scenario Context
Initial Access	TA0001	Phishing email leading to network infiltration; exploitation of unpatched software vulnerabilities on the HMI/SCADA systems
Execution	TA0002	Ransomware payload execution encrypting SCADA control system data and HMI displays
Impact	TA0040	Data encryption for ransom; disruption of water treatment process control; potential public health impact through service interruption

Implied TTPs Across All Trials

While only A12 explicitly triggered TTP reporting, the scenario parameters and transcript narratives across all 20 trials consistently described the following adversary behaviors that map to MITRE ATT&CK:

Tactic	ID	Technique	Evidence from Transcripts
Initial Access	TA0001	Phishing (T1566)	A2, A4 AARs identified phishing as the suspected attack vector; A16 recommended phishing simulations
Execution	TA0002	User Execution (T1204)	Ransomware deployed after user interaction with phishing payload
Persistence	TA0003	Valid Accounts (T1078)	Weak network segmentation allowed lateral movement; unpatched systems provided persistent access
Lateral Movement	TA0008	Exploitation of Remote Services (T1210)	A1 transcript: "weak network segmentation allowing ransomware to spread among critical systems"
Impact	TA0040	Data Encrypted for Impact (T1486)	All 20 transcripts: ransomware encrypts HMI/SCADA data; ransom demanded in Bitcoin (50 BTC in A1)
Impact	TA0040	Service Stop (T1489)	HMI becomes non-functional; automated water treatment process disrupted across all trials

Adversary Agent Behaviors

AutoTableTop™ 2.0 includes an adversary agent that escalates the scenario in parallel with participant responses. Across all 20 transcripts, the adversary agent consistently executed the following actions:

- **Communication system disruption:** Targeting emergency communication infrastructure to delay coordination among response agencies.
- **Social media misinformation campaigns:** Spreading false information about crisis severity, evacuation routes, shelter locations, and government response effectiveness.
- **Supply chain interference:** Obstructing transportation routes and manipulating logistics to delay delivery of essential supplies.
- **Resource misallocation:** Generating false emergency calls and reports to redirect resources away from actual crisis zones.

These adversary behaviors map to a sophisticated threat actor profile combining cyber-physical attack capabilities with information operations, consistent with nation-state or advanced persistent threat (APT) actors targeting critical infrastructure.

Input Phrasing Analysis

Inputs That Consistently Score 7 - 8

Phase 1: Initial Response

Best phrasing: "*Notify the proper authorities of the incident. Switch the process operated by the HMI to manual operations and hand the operations off to a human operator for now.*"

Used in all 20 trials, scored 7–8 every time. Every AAR cited it as a successful action. Combines notification + manual switchover in a single decisive statement.

Phase 2: DFIR Engagement

Best phrasing (scores 8): "*Does the water facility have an ongoing contract with a DFIR company or other means of access to knowledgeable cybersecurity incident response? Request assistance and hand them the information and access necessary to assist in removing the threat.*"

The question-form consistently outperformed the declarative form (8 vs. 7) because the judge rewards situational awareness. The AARs reinforced this, A7 and A18 both recommended establishing DFIR contracts as safety recommendations.

Phase 3: Stakeholder Communication

Best phrasing (scores 8): "*Prepare a public statement and contact any legal stakeholders to provide a proper response for the federal authorities.*"

Combines public, legal, and regulatory communication in one action. AARs from A3, A9, A14, A15 scored this 7 to 8. The AARs' recurring improvement area around public communication validates the importance of including this step explicitly.

Phase 4: Resolution Monitoring

Best phrasing (scores 7): "*Continue until the cybersecurity team has finished their efforts on removing the ransomware infection, then when the ransomware infection has been removed, reintroduce the HMI and switch back to automatic operations.*"

Forward-planning compound statement. The AARs' consistent flagging of passive waiting as an improvement area validates this approach, pre-committing to the post-resolution action scores higher than sequential waiting.

Phase 5: Debrief & Closure

Best phrasing (scores 7–8): "*Monitor the ongoing situation and prepare a debrief and lessons learned documentation to discover the causes and potential solutions to similar incidents.*"

Combines vigilance with forward-looking analysis. AARs from A2, A5, and A11 scored this action 8.

Inputs That Score Poorly (0 to 4)

The following patterns were flagged as improvement areas in the AARs and scored 0–4 in the transcripts:

- "*OK.*" — Scored 0 in every instance. AARs flagged as "minimal response acknowledging the situation."
- "*Await resolution*" / "*Continue supporting the team.*" — Scored 4–5. AARs flagged as "passive waiting without active engagement."
- "*Provide the confirmation when ready.*" — Scored 2. No proactive value.
- "*OK, notify me when the team is done.*" — Scored 3. AARs noted "ensuring active engagement and action follow-ups" as the improvement area.

Key Findings

Optimal Turn Count

The ideal number of facilitator inputs is 4 to 6. Trials with fewer than 4 inputs resolved too quickly and missed stakeholder communication steps. Trials with more than 8 inputs introduced filler turns that lowered averages. The AARs confirmed this: trials with 4 to 7 inputs received higher containment ratings than those with 10+ inputs.

Five-Phase Decision Framework

Every high-scoring trial and every AAR-rated "*Successful*" containment followed this framework:

1. Immediate Response: Notify authorities + switch to manual operations
2. Expert Engagement: Mobilize DFIR/cybersecurity resources with clear handoff
3. Stakeholder Communication: Public statement + legal/regulatory notification
4. Monitored Resolution: Support DFIR through completion with pre-planned transition
5. Post-Incident Review: Lessons learned documentation + structured debrief

AAR Quality Correlation

A strong correlation exists between transcript average score and AAR quality indicators. Trials scoring 7.0+ received AAR ratings of "*High*" for multi-agency coordination and "*Successful*" for overall containment. Trials scoring below 6.0 received "*Fair*" or "*Moderate*" ratings and had significantly more improvement areas flagged.

TTP Mapping Threshold

The MITRE ATT&CK tactic mapping appears to require sufficient technical narrative depth to trigger. Only A12, which included detailed restoration sequencing and explicit discussion of the ransomware's behavior, produced TTP output. For demo purposes, the implied TTPs reconstructed from scenario parameters provide equivalent value and should be presented alongside the AAR.

The Perfect Demo Script

Based on the integrated analysis of all 20 transcripts and 20 AARs, the following 6-input script is optimized for the highest possible scores, the richest AAR output, and complete demonstration of AutoTableTop™ 2.0's crisis management exercise capabilities.

SCENARIO SETUP

The system will present the scenario introduction: a ransomware note is discovered on the HMI screen at a Texas water treatment facility. The facilitator reads the briefing and prompts for the first participant action.

INPUT 1: Initial Response

Expected Score: 7 to 8

Phase: Detection & Containment

Script: *"Notify the proper authorities of the incident. Switch the process operated by the HMI to manual operations and hand the operations off to a human operator for now."*

Why this works: Scored 7 to 8 in all 20 trials. Cited as a successful action in every AAR. Combines two critical first-response actions: activating the notification chain (law enforcement, CISA, regulatory bodies) and ensuring operational continuity through manual operations.

AAR validation: All 20 AARs listed this as a successful action. A16's AAR noted "*initial actions taken within 30 minutes of incident discovery.*" A19 scored this an 8 for "prioritizing life safety and enabling prompt response."

INPUT 2: DFIR Engagement

Expected Score: 8

Phase: Expert Mobilization

Script: *"Does the water facility have an ongoing contract with a DFIR company or other means of access to knowledgeable cybersecurity incident response? Request assistance and hand them the information and access necessary to assist in removing the threat."*

Why this works: The question-form consistently scored 8. The judge recognizes situational awareness—assessing available resources before committing to action.

AAR validation: A2 scored this 8 for "proactive threat containment and expert involvement." A18's primary safety recommendation was "Establish a DFIR contract for rapid response." A7's AAR flagged the absence of a DFIR contract as an improvement area.

INPUT 3: Stakeholder Communication & Legal Preparation

Expected Score: 7 to 8

Phase: Communication & Compliance

Script: *"Prepare a public statement and contact any legal stakeholders to provide a proper response for the federal authorities."*

Why this works: Addresses public, legal, and regulatory communication vectors in a single compound action. The AARs' most frequently cited improvement area, public communication, validates the importance of this step.

AAR validation: A9 scored this 8 for engaging legal stakeholders. A12 scored 8 for "*contacting stakeholders and preparing a public statement to manage crisis communications.*" Six AARs recommended "robust public communication strategies" as a safety recommendation.

INPUT 4: Monitored Resolution with Forward Planning

Expected Score: 6 to 7

Phase: Containment & Recovery

Script: *"Continue until the cybersecurity team has finished their efforts on removing the ransomware infection, then when the ransomware infection has been removed, reintroduce the HMI and switch back to automatic operations."*

Why this works: The resolution bridge, pre-commits to post-resolution action. Scores 6 to 7 because monitoring is inherently less active, but the compound structure outperforms sequential inputs.

AAR validation: AARs from A6, A8, A12, A20 flagged transition planning as needing careful evaluation. This phrasing explicitly ties HMI reintroduction to DFIR confirmation, addressing that concern. A19 scored this 7 for "*ensuring operational security.*"

INPUT 5: Post-Incident Review

Expected Score: 7 to 8

Phase: Lessons Learned & Closure

Script: *"Monitor the ongoing situation and prepare a debrief and lessons learned documentation to discover the causes and potential solutions to similar incidents."*

Why this works: Combines ongoing vigilance with forward-looking analysis. The specific language about "*causes and potential solutions*" adds substance.

AAR validation: A5 scored this 8: "*Preparation of debrief and lessons learned document.*" A11 scored 8 for "*conducting a thorough post-action analysis.*" A19 scored 7 for "*fostering continuous improvement.*"

INPUT 6: Exercise Closure

Expected Score: 7–8

Phase: Wrap-Up

Script: *"It seems like the incident has been resolved. Is there any other steps that are worth taking within this exercise?"*

Why this works: Scored 8 in A3. The open check demonstrates thoroughness, ensuring nothing was missed before concluding.

AAR validation: A11's AAR scored a similar closure prompt 8 for "*thorough post-action analysis.*" A14 scored 8 for explicit closure framing. Clean closure produces richer AAR output with more complete crisis response metrics.

EXPECTED OUTCOME

Projected Average Score: 7.2 to 7.5

Projected Turn Count: 6 inputs

Projected Score Distribution: 7, 8, 8, 6, 8, 8 (based on historical phrasing performance)

Projected AAR Quality: Multi-Agency Coordination: High | Overall Containment: Successful

Projected Demo Duration: approximately 20 minutes of active interaction

This script demonstrates AutoTableTop™ 2.0's full exercise lifecycle: realistic scenario delivery, adaptive narrative progression, per-turn scoring, adversary agent escalation, and comprehensive after-action reporting with actionable recommendations.

Quick Reference Card

Condensed demo script for live presentations. Print or display this single page for the presenter.

AUTOTABLETOP™ 2.0 — DEMO SCRIPT (WATER/WASTEWATER RANSOMWARE)

1. INITIAL RESPONSE: “Notify the proper authorities of the incident. Switch the process operated by the HMI to manual operations and hand the operations off to a human operator for now.”

→ *Wait for facilitator narrative response*

2. DFIR ENGAGEMENT: “Does the water facility have an ongoing contract with a DFIR company or other means of access to knowledgeable cybersecurity incident response? Request assistance and hand them the information and access necessary to assist in removing the threat.”

→ *Wait for facilitator narrative response*

3. STAKEHOLDER COMMS: “Prepare a public statement and contact any legal stakeholders to provide a proper response for the federal authorities.”

→ *Wait for facilitator narrative response*

4. RESOLUTION: “Continue until the cybersecurity team has finished their efforts on removing the ransomware infection, then when the ransomware infection has been removed, reintroduce the HMI and switch back to automatic operations.”

→ *Wait for facilitator narrative response*

5. POST-INCIDENT: “Monitor the ongoing situation and prepare a debrief and lessons learned documentation to discover the causes and potential solutions to similar incidents.”

→ *Wait for facilitator narrative response*

6. CLOSURE: “It seems like the incident has been resolved. Is there any other steps that are worth taking within this exercise?”

→ *System concludes exercise and generates After-Action Report*

Target: 6 inputs | ~15–20 min | Projected avg score: 7.2 to 7.5 | AAR: High coordination, Successful containment

Appendix A: Scenario Parameters

The following JSON parameters were used to generate the scenario for all 20 trials:

Parameter	Value
Threat Scenario	Ransomware note on HMI screen at water treatment facility
Company	Texas Water Treatment Facility
Department	Water Plant Operators
Industry	Water & Waste Water
Exercise Type	Crisis Management
Difficulty	Beginner
Technical Level	Basic
Verbosity	Standard
Complexity	Beginner
Win Condition	Facility operations return to normal
Lose Condition	Facility can no longer operate normally or significant loss of life

Appendix B: AAR TTP Demo Quick Reference

Since only 1 of 20 trials triggered automatic MITRE ATT&CK TTP mapping in the AAR, presenters should be prepared to discuss the implied TTPs when demonstrating the After-Action Report. The following reference card covers the relevant TTPs for this scenario:

Tactic	ID	Technique	Presenter Talking Point
Initial Access	T1566	Phishing	The ransomware entered through a phishing email exploiting unpatched systems
Execution	T1204	User Execution	An operator likely clicked a malicious link or attachment
Lateral Movement	T1210	Exploit Remote Svcs	Weak network segmentation allowed the ransomware to spread to SCADA/HMI
Impact	T1486	Data Encrypted	HMI and control system data encrypted; 50 BTC ransom demanded
Impact	T1489	Service Stop	Water treatment automation disrupted, forcing manual operations

Appendix C: industry competition

This scenario was developed for use as the cybersecurity demonstration event at a 2026 water industry competition, hosted by a state water environment association. The competition took place in Spring 2026.

Event Overview

This water industry competition was a multi-event competition recognizing excellence and professionalism in wastewater treatment operations, maintenance, laboratory, safety, and collection systems. The 2026 competition featured multiple teams across two divisions competing in five primary scored events (Pump Maintenance, Collection System, Safety, Process Control, and Laboratory) plus six demonstration events.

The cybersecurity demonstration event was one of the demonstration events, described as "*a desktop cyber security crisis simulator and response planning.*" AutoTableTop™ 2.0 was the platform powering this demonstration event. The event coordinator is the event coordinator from a participating river authority.

Participant Profile

The competition participants were water and wastewater operations professionals, including plant operators, maintenance technicians, laboratory personnel, and collection systems staff. These individuals **possess deep expertise** in plant operations but typically limited cybersecurity background. This profile directly informed the scenario parameters:

- Exercise Type: Crisis Management (not technical incident response)
- Difficulty Level: Beginner
- Technical Level: Basic
- Verbosity Level: Concise
- Participants: "*Operators and not cyber focused*"

This calibration ensures the exercise meets participants where they are - testing crisis management decision-making and operational continuity skills rather than technical cybersecurity knowledge.

Event Structure Relevance

The cybersecurity demonstration event was run with teams cycling through in approximately 30-minute time slots. This constraint validates the optimized 6-input demo script's projected 15 to 20 minute runtime, and it fit within the event's scheduling framework while leaving time for team transitions and brief discussion.

Appendix D: Expert Review, Randy Petersen

Reviewer Background

Randy (William) Petersen is the SCADA Superintendent at San Jacinto River Authority (SJRA), where he has spent over thirteen years focused on operational technology with a strong emphasis on networking and cybersecurity. He is responsible for the collective network and server infrastructure, programming, and cybersecurity of SJRA's industrial control systems—an environment spanning over 1,200 network devices including PLCs, HMIs, SCADA servers, and field instruments across multiple divisions.

Petersen is a recognized authority in the OT cybersecurity community. He is a co-founder of the SCADA User Group, serves on the TAWWA Resiliency and Cybersecurity Committee, holds ISC2 certification, and is a frequent speaker at industry conferences including OT.SEC.CON., HOU.SEC.CON., and CYBR.SEC.CON. He has published on topics including network asset visibility, OT cybersecurity program development, and the availability-first mindset that distinguishes OT security from IT security. His NACWA publication "*10 Practical Steps to Reduce SCADA Cybersecurity Risk*" is widely referenced in the water sector.

Petersen provided the original scenario configuration and conducted an independent evaluation of AutoTableTop™ 2.0 using the event settings.

Scenario Configuration Review

Petersen's configuration for the event differed from the 20-trial study parameters in several notable ways:

Parameter	20-Trial Study	Petersen Config
Company Name	Texas Water Treatment Facility	***
Company Info	"***"	(Left blank)
Objectives	Detailed detection/containment/recovery	"*Continuous delivery of water*"
Participants	Water plant operators with ops knowledge	"*Operators and not cyber focused*"
Verbosity	Standard	Concise
Other Fields	Most set to "Surprise me"	Most left blank

Key observation: Petersen's configuration is deliberately minimalist, reflecting a practitioner's approach. By leaving most fields blank rather than using "*surprise me,*" he tested whether AutoTableTop™ 2.0 could generate a coherent and useful exercise with minimal input, a critical capability for real-world deployment where busy operators may not have time to fill out detailed configuration forms. The single-line exercise objective ("*Continuous delivery of water*") is notably operator-centric, focusing on the operational outcome rather than cybersecurity process.

Petersen's Exercise Transcript Analysis

Petersen conducted a live exercise run on February 25, 2026, using his event configuration. His transcript reveals a distinctly operator-focused response pattern:

Input 1: "*Read note and immediately call supervisor*"

Score: 6/10

This response reflects authentic operator behavior, a water plant operator's first instinct is to read the note and escalate to their supervisor, not to initiate an incident response playbook. The lower score (6 vs. the study's typical 7 to 8 for Turn 1) is attributable to the absence of a concurrent operational action (the study's optimized script adds "*switch to manual operations*" alongside notification). However, this is exactly the response an actual plant operator would give, validating the scenario's accessibility to the target audience.

Input 2: "*Team decides that since there is no visibility on the HMI to start doing physical rounds*"

Score: 6/10

This is the most operationally authentic response in the entire study. No trial in the 20-trial study produced this action. Physical rounds - walking the plant to visually inspect equipment, check gauges, verify chemical feed systems, and ensure pumps are operating - is the fundamental fallback procedure for water plant operators when automated monitoring fails. It demonstrates the exact skill set the exercise is designed to test: operational continuity under degraded conditions.

The judge scored this 6 rather than 7 to 8 because the action lacks explicit authority notification and cybersecurity engagement. From a crisis management scoring perspective, this is a valid assessment. But from an operational authenticity perspective, this response validates that the scenario successfully engages operators at their level of expertise.

Input 3: "*End exercise and generate report*"

Petersen concluded the exercise after two inputs to evaluate the AAR generation capability. The exercise was completed with only two substantive actions, which limited the depth of the AAR analysis but demonstrated the platform's ability to generate a complete report from minimal input.

Petersen's AAR Results

The After-Action Report generated from Petersen's abbreviated exercise:

- Overall Crisis Containment: Partial, with room for improvement
- Multi-Agency Coordination: Moderate
- Total Casualties: 0 (Prevented)
- Public Communication Effectiveness: Limited due to network issues
- Key Improvement Areas: Structured communication during physical rounds; contingency plans for communication disruptions

- Key Operational Learning: "*Physical rounds can compensate for technological failures but require structured communication plans*"

The AAR correctly identified the exercise's limitations and generated actionable recommendations despite receiving only two participant inputs, demonstrating robust report generation from sparse data.

Recommended Event Settings

Based on Petersen's configuration, the 20-trial study findings, and the event's scheduling constraints, the recommended settings for the water industry operations competition cybersecurity event are:

Parameter	Recommended Value
Company Name	***
Exercise Objectives	Continuous delivery of water
Participants	Participants will be operators and not cyber focused
Exercise Type	Crisis Management
Difficulty	Beginner
Verbosity	Concise
Technical Level	Basic
Complexity	Beginner
All other fields	(Leave blank — AutoTableTop™ fills reasonable defaults)

