

THE HUMAN CEILING

The cognitive limits of live video monitoring in enterprise security operations, and what they cost.

Prepared by Jody Russell, AI Evangelist and Solutions Engineer

WHAT THIS WHITE PAPER DOES NOT CLAIM

To keep the argument honest, it is worth stating plainly what follows from the evidence and what does not.

This white paper does not claim that security operators are careless or unprofessional; that AI is infallible or error-free; that every incident is preventable; that cameras alone prevent harm; or that healthcare alarm fatigue is identical to security alarm fatigue. It does not claim that headcount has no value. It makes one structural claim: the live-monitoring model cannot scale without a machine-supported detection layer. Everything else in these pages is in service of that single point.

EXECUTIVE SUMMARY

Enterprise surveillance has scaled past the biological limits of the people asked to watch it. This is a measurement problem with a measurable cost.

Over the last two decades, the enterprise camera estate has grown by an order of magnitude while the human visual system has not changed at all. The result is a structural mismatch: organizations now operate camera fleets whose volume of live imagery exceeds what any realistic number of operators can meaningfully observe. The consequence is not occasional human error. It is a predictable, well-documented decline in detection that begins within minutes of a shift and compounds across every additional screen, every false alarm, and every overnight hour.

This white paper assembles the peer-reviewed and government-sourced evidence behind that claim and translates it into operational and financial terms for executive decision-makers. Four findings anchor the analysis.

EXECUTIVE SUMMARY

First, sustained attention has a short half-life. The vigilance decrement, the decline in detection of rare events over time on task, is one of the most replicated findings in experimental psychology, first quantified by Mackworth in 1948 and confirmed continuously since. Measurable decline typically sets in within the first 15 to 30 minutes and is largely involuntary; no amount of diligence repeals it.

Second, detection erodes as feeds multiply. When operators monitor many camera feeds at once, correct detection of staged events falls measurably, from roughly 84% on nine monitors to about 64% on sixteen in the classic experiment, an effect reproduced in modern closed-circuit television (CCTV) research. The same cognitive science explains why even trained experts miss salient, unexpected events in plain view: inattention blindness is a feature of human attention, not a lapse of professionalism.

Third, alarm fatigue is a rigorously documented phenomenon with documented safety consequences, including deaths. The strongest empirical base is in medicine, where 85–99% of monitor alarms require no intervention, a single hospital can log millions of alarms per month, and regulators have tied dozens of patient deaths to alarm desensitization. The mechanism, humans tuning out high-volume, low-yield alerts, has a clear parallel in a Security Operations Center (SOC), where 94–98% of police alarm responses are to false alarms. We treat the medical evidence as a rigorous analogy, not as direct proof of SOC behavior; the security industry has far less direct empirical research of its own.

Fourth, the labor model amplifies every weakness above. A workforce defined by high turnover, modest wages, and around-the-clock shift work is asked to perform a task that degrades with fatigue. Peer-reviewed research equates 17 hours awake with a 0.05% blood-alcohol concentration (BAC) and 24 hours with roughly 0.10%, above the legal driving limit. Staffing a single post around the clock requires more than four full-time employees (FTEs) before relief factors; doing it well, at fleet scale, is economically out of reach for most organizations.

Taken together, these findings reframe a familiar assumption. The problem with enterprise video is not that organizations have too few cameras or too little footage. It is that the live-monitoring model relies on a human capability that does not scale and was never designed for the load now placed on it. Real-world failures across corporate, critical-infrastructure, and transit settings illustrate the gap, and the academic record is specific about the consequence: passively recorded video shows small, non-significant preventive effects, whereas actively monitored video produces statistically significant crime reductions. The strategic question for executives is therefore not whether to watch more, but how to ensure that what matters is actually seen: at scale, consistently, and in time to act.

The closing sections translate this into a practical model in which artificial intelligence performs the continuous, fatigue-free evaluation that humans cannot sustain, while skilled officers are elevated to verification, judgment, escalation, and response. This is the design principle behind Ambient.ai's Agentic Physical Security platform, and it is grounded directly in the evidence presented here.

01 · THE SCALE PROBLEM

The volume of live imagery in a modern enterprise has outgrown the number of human eyes any budget can put in front of it.

The defining fact of contemporary physical security is abundance. Industry market research counted roughly 770 million surveillance cameras installed worldwide in 2019; the global installed base passed one billion in 2021 and, on current trajectories, is projected to approach two billion by 2027. Within a single large enterprise (a global corporate campus network, a utility's generation and substation fleet, a transit authority's stations and tunnels), camera counts routinely run from the hundreds into the tens of thousands.

Surveillance has outgrown human capacity to watch it

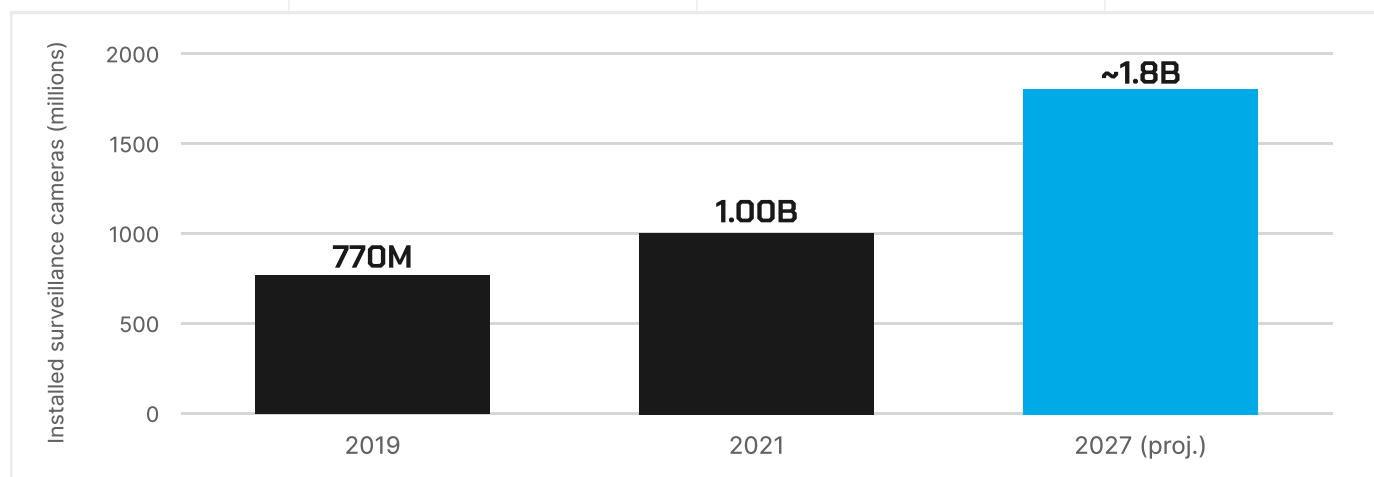


Figure 1. Global installed base of surveillance cameras (770M in 2019; one billion in 2021; ~1.8B projected for 2027). Source: IHS Markit / Omdia, reported by CNBC.

The human side of the equation tells the opposite story. A control-room operator does not gain capacity as cameras are added; capacity is fixed by biology. As the feed-to-operator ratio rises, each individual feed receives a smaller share of genuine attention until, past a threshold, additional cameras contribute almost nothing to real-time awareness. They become a forensic archive: useful after an incident, largely inert before one.

A field study of a municipal CCTV expansion captured this directly. When one city grew its system from 14 to 146 cameras, researchers documented up to a 47% reduction in proactive surveillance activity, and the camera-to-operator ratio was adjusted downward from 73:1 to 33:1 in an attempt to compensate, an explicit acknowledgement that the system had outgrown its watchers. The same review notes how rarely live monitoring produces an intervention at all: in one widely cited evaluation, operators averaged roughly one arrest per 967 hours of monitoring.

This is the backdrop against which every other finding should be read. The cognitive limits described next are not new. They have been measured for three-quarters of a century. What is new is the scale at which organizations now collide with them.



02 · THE SCIENCE OF ATTENTION

Sustained vigilance for rare events is among the hardest tasks in cognitive psychology, and surveillance monitoring is its purest real-world form.

Watching monitors for something that almost never happens is, in human-factors terms, a vigilance task: the observer must sustain attention over long periods to detect infrequent, unpredictable, low-probability signals. Decades of research establish that human performance on such tasks deteriorates rapidly and largely involuntarily. The phenomenon is called the vigilance decrement, and it is among the most robust findings in experimental psychology.

The foundational study dates to 1948, when Norman Mackworth asked observers to monitor a clock-like pointer and report rare irregular jumps over a two-hour watch. Detection of the rare signals declined over the vigil, with the decline becoming pronounced after roughly the first half hour, the first laboratory quantification of attention breaking down under sustained monitoring. A 2025 retrospective marking the decrement's first 75 years confirms that the effect has survived every subsequent attempt to design it away, across modalities, incentives, and observer populations. Modern reviews place the onset of measurable decline within the first 15 to 30 minutes of continuous monitoring.

The vigilance decrement: monitoring performance decays within minutes

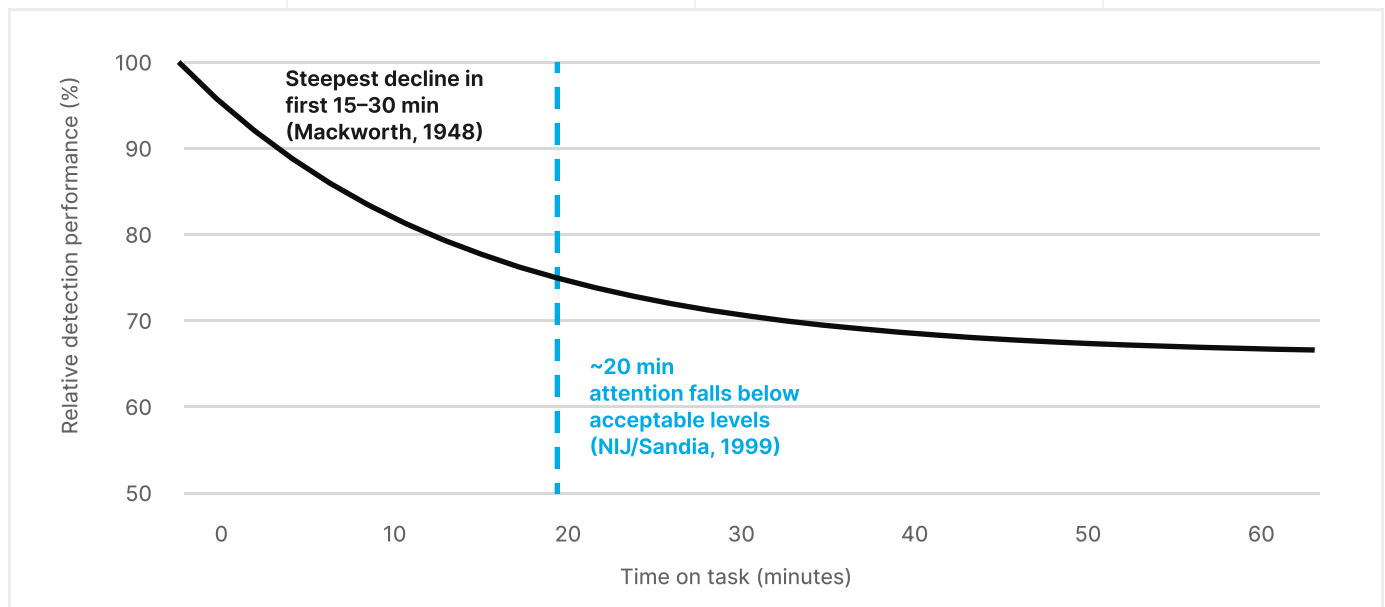


Figure 2. The vigilance decrement: relative detection performance declines steeply early in a monitoring period, then more gradually. Illustrative synthesis of the vigilance literature (Mackworth, 1948; National Institute of Justice / Sandia, 1999); shape is representative, not a single dataset.

What the government guidance actually says, and does not

The most frequently quoted statement in the security industry on this subject comes from a 1999 guide prepared by Mary W. Green of Sandia National Laboratories for the U.S. National Institute of Justice (NIJ). The widely circulated rendering reads:

**“After only 20 minutes of watching and evaluating monitor screens,
the attention of most individuals has degenerated to well below acceptable levels.
Monitoring video screens is both boring and mesmerizing.”**

A 2002 U.S. Government Accountability Office (GAO) testimony on federal building security restates the same finding nearly verbatim: “Because watching camera screens is both boring and mesmerizing, the attention of most individuals has degenerated to well below acceptable levels after only 20 minutes of viewing.” It is important to be precise about what these sources do and do not establish. They assert, on the basis of practitioner and laboratory experience, that attention degrades to unacceptable levels in roughly twenty minutes. They do not attach a specific miss-rate percentage to that statement, and any retelling that cites a precise “45% after 12 minutes / 95% after 22 minutes” figure is repeating a number with no traceable source. We reject that statistic and rely only on what the primary documents actually say.

The practical implication is structural. Because the decrement is involuntary and fast-onset, it cannot be solved by hiring more conscientious people or by exhortation. It is a property of the human attentional system performing a task it is poorly suited to: exactly the task that 24/7 live monitoring demands.

03 · THE MULTIPLEXING PROBLEM

Detection does not merely decline over time. It declines with every additional feed competing for the same fixed pool of attention.

If a single feed already taxes sustained attention, multiple simultaneous feeds compound the problem. The capacity of human attention to track concurrent streams of information is sharply bounded, and surveillance monitoring runs straight into that bound.

Detection falls as feeds rise

The foundational experiment is Tickner and Poulton's 1973 study in Ergonomics, which measured how accurately observers detected staged suspicious events while monitoring 4, 9, or 16 television pictures at once. Detection held near 83–84% at four and nine monitors but fell to roughly 64% at sixteen. That result is the empirical origin of the long-standing operational guideline that a single operator should watch no more than about sixteen feeds, and far fewer when scenes contain heavy movement. Modern peer-reviewed CCTV research reproduces the underlying effect: in one field-realistic study, operators detected only about half of target behaviours, and vigilance degraded with both exposure and time on task.



Detections collapse as feeds per operator rise

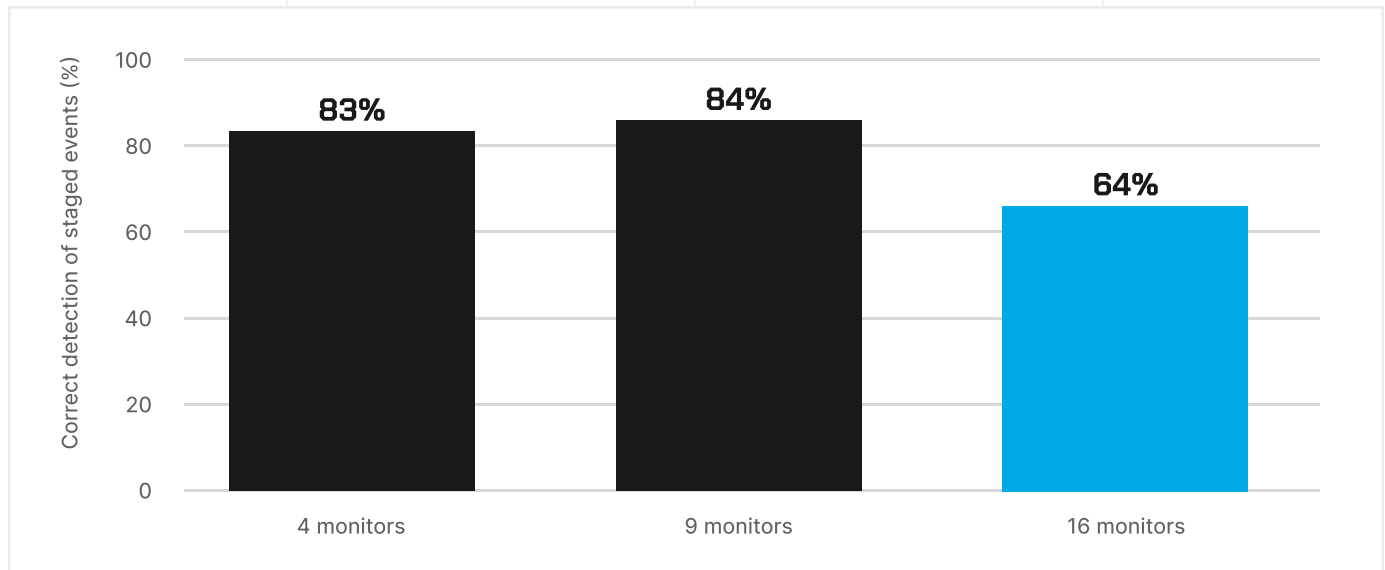


Figure 3. Correct detection of staged events by number of simultaneously monitored feeds. Source: Tickner & Poulton, *Ergonomics* 16(4), 1973. Detection is roughly stable to nine feeds, then drops sharply at sixteen.

The working-memory ceiling

Why detection collapses is explained by the architecture of attention itself. The focus of attention holds only a handful of items at once. While Miller's classic "seven, plus or minus two" estimate is the famous figure, the more defensible modern limit is Cowan's: the focus of attention reliably holds about four chunks of information (three to five) in normal adults. A wall of two dozen monitors does not expand this limit; it simply guarantees that most of what is on screen at any instant falls outside it.

Inattention blindness: looking is not seeing

Even within the few feeds an operator is actively attending, a salient event can pass unnoticed. In the canonical demonstration, Simons and Chabris asked observers to count basketball passes while a person in a gorilla suit walked through the scene; about half failed to notice the gorilla at all. When attention is loaded by a task, the visual system can fail to register even large, unexpected, fully visible events. Expertise does not confer immunity: a 2013 study found that 83% of trained radiologists missed a gorilla image inserted into a computed tomography (CT) scan they were examining.

Together these three results (feed-count degradation, the working-memory ceiling, and inattention blindness) explain an uncomfortable operational reality: a control room can be fully staffed, fully attentive, and fully professional, and still miss the one event that matters. The failure is designed into the task, not into the people.



04 · ALARM FATIGUE: THE COST OF CRYING WOLF

When almost every alert is false, humans rationally learn to discount alerts. The most rigorous evidence comes from medicine, and it includes deaths.

Automation was supposed to rescue the overwhelmed operator: instead of watching everything, watch the alarms. But alarms only help if they are trustworthy, and in practice they are overwhelmingly not. The predictable human response to a high volume of low-yield alerts is desensitization, or alarm fatigue, in which genuine alarms are missed, silenced, or disabled because the great majority before them were false.

The healthcare evidence base

The phenomenon is best documented in hospitals, where it has been studied with a rigor the security industry has not yet matched. The Joint Commission, the principal U.S. hospital accreditor, issued a Sentinel Event Alert in 2013 estimating that 85 to 99 percent of alarm signals require no clinical intervention, with a single patient generating up to several hundred alarms per day. A peer-reviewed observational study at the University of California, San Francisco (UCSF) crystallized the scale: across five intensive-care units (ICUs) over 31 days, monitors generated 2,558,760 alarms, and 88.8% of the arrhythmia alarms were false positives.

Alarm fatigue is an evidence base, not a metaphor

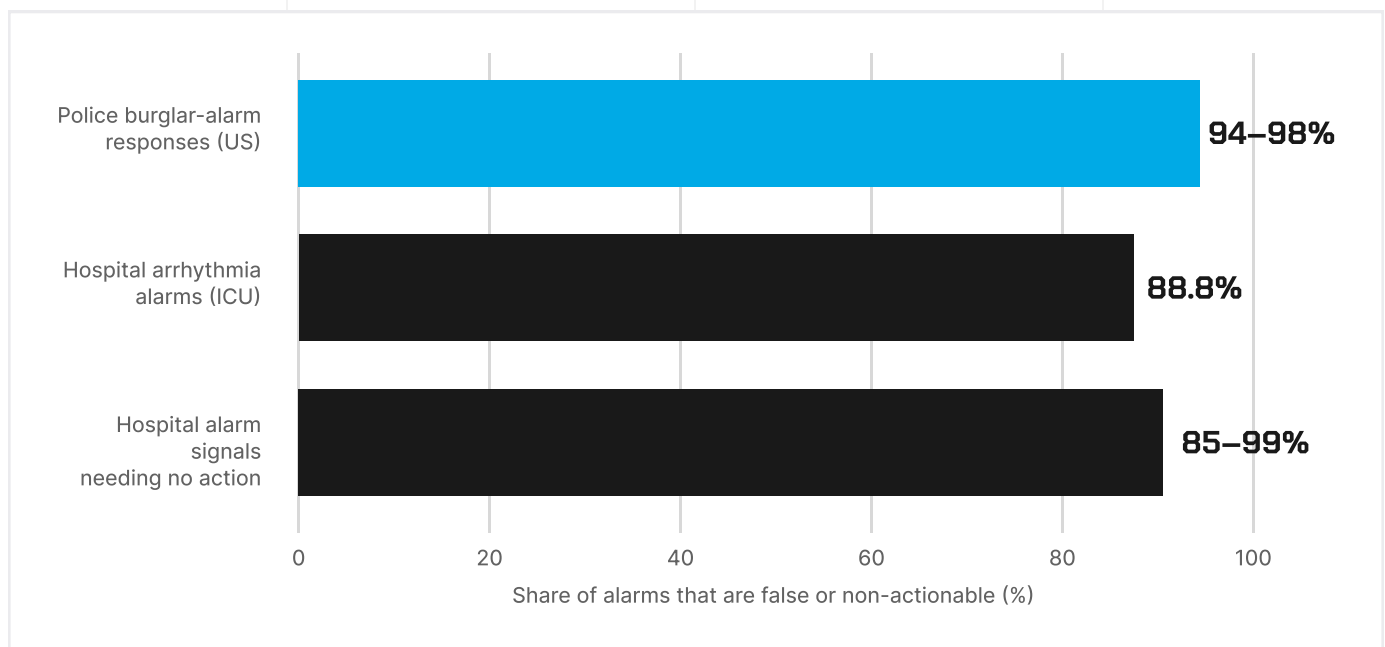


Figure 4. The share of alarms that are false or require no action, across policing and healthcare. Sources: U.S. Department of Justice, Office of Community Oriented Policing Services (Sampson, 2007); Drew et al., PLOS ONE, 2014; The Joint Commission Sentinel Event Alert #50, 2013.

This is not a quality-of-life nuisance; it has documented safety consequences, including deaths. The Joint Commission identified 98 alarm-related events reported between January 2009 and June 2012, of which 80 resulted in patient death and 13 in permanent loss of function, with alarm fatigue cited as the most common contributing factor. Separately, the Manufacturer and User Facility Device Experience (MAUDE) database maintained by the U.S. Food and Drug Administration (FDA) recorded 566 alarm-related deaths over the five-year window from January 2005 to June 2010. Independent patient-safety research ranked clinical-alarm hazards the number-one health-technology hazard for four consecutive years (the lists for 2012 through 2015).

The same mechanism in security, by analogy

The cognitive mechanism is domain-independent: a security operator flooded with false alerts is subject to the same desensitization a nurse is. We present this as a strong analogy rather than direct proof, because the security industry has not produced the controlled studies that medicine has. What the security industry does have is striking false-alarm data: a U.S. Department of Justice (DOJ) problem-oriented policing guide documents that 94 to 98 percent of police alarm responses are to false alarms, with individual jurisdictions reporting valid-alarm rates as low as 0.3%. The downstream cost is concrete: roughly 36 million alarm activations a year in the United States, at an estimated \$1.8 billion in annual response cost.

05 · THE HUMAN OPERATOR UNDER LOAD

The task degrades with fatigue. The staffing model runs on fatigue. The two collide every night shift.

Everything to this point assumes a rested, alert operator. Real security operations do not run on rested operators. They run around the clock, on nights and rotating shifts, often with long hours, precisely the conditions the human-factors literature shows degrade attention and judgment most severely.

Fatigue is impairment, measured against a familiar yardstick

The landmark quantification comes from Dawson and Reid, published in *Nature* in 1997. They found that after 17 hours of sustained wakefulness, cognitive performance fell to a level equivalent to a BAC of 0.05%, and after 24 hours to roughly 0.10%, above the 0.08% legal driving limit in most U.S. states. An independent study by Williamson and Feyer reproduced the result: 17 to 19 hours without sleep impaired performance to roughly a 0.05% BAC equivalent, with response speeds up to 50% slower, and passive vigilance tasks, exactly the kind surveillance monitoring demands, degrading earliest.



A fatigued operator is an impaired operator

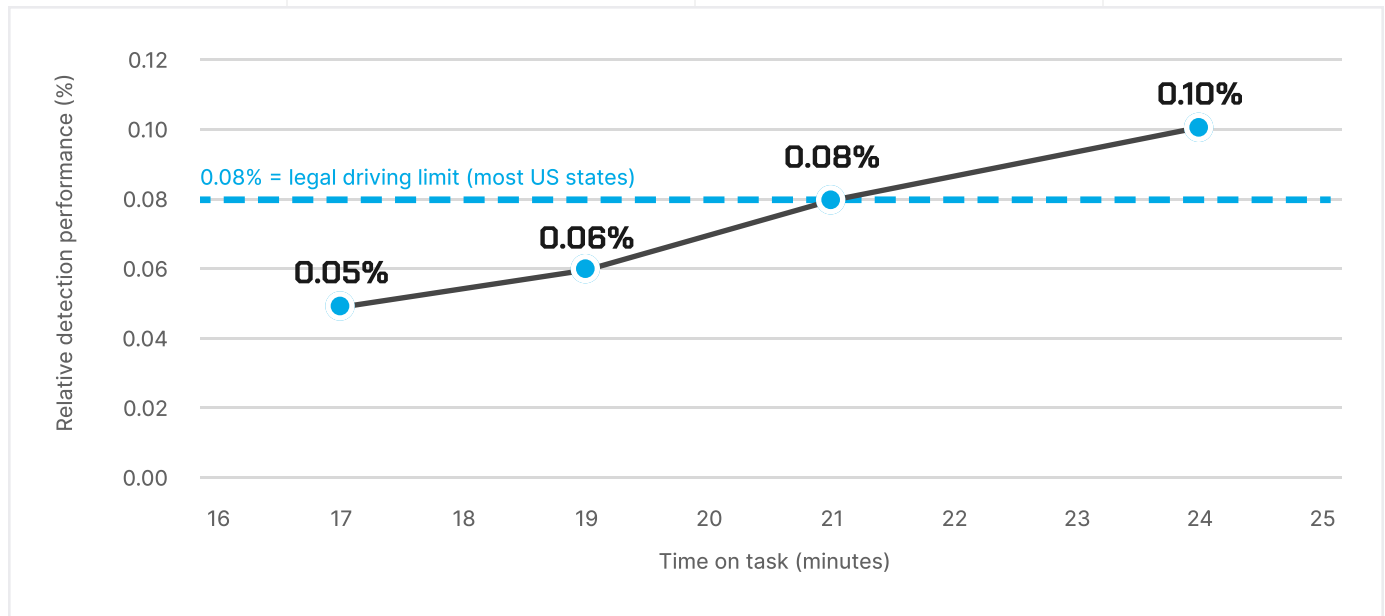


Figure 5. Cognitive impairment from sustained wakefulness, expressed as equivalent blood-alcohol concentration. Source: Dawson & Reid, Nature 388:235, 1997; corroborated by Williamson & Feyer, 2000. The dashed line marks the 0.08% legal driving limit.

Shift design compounds the risk

A large modeling study by Folkard and Lombardi quantified how injury and incident risk rises with shift type and length: relative to a morning shift, risk is about 27.9% higher on night shifts, rising to roughly 36% higher by the fourth consecutive night. Risk also climbs with shift length: about 13% higher on 10-hour shifts and 27.5% higher on 12-hour shifts versus 8-hour shifts.

Shift design compounds cognitive risk

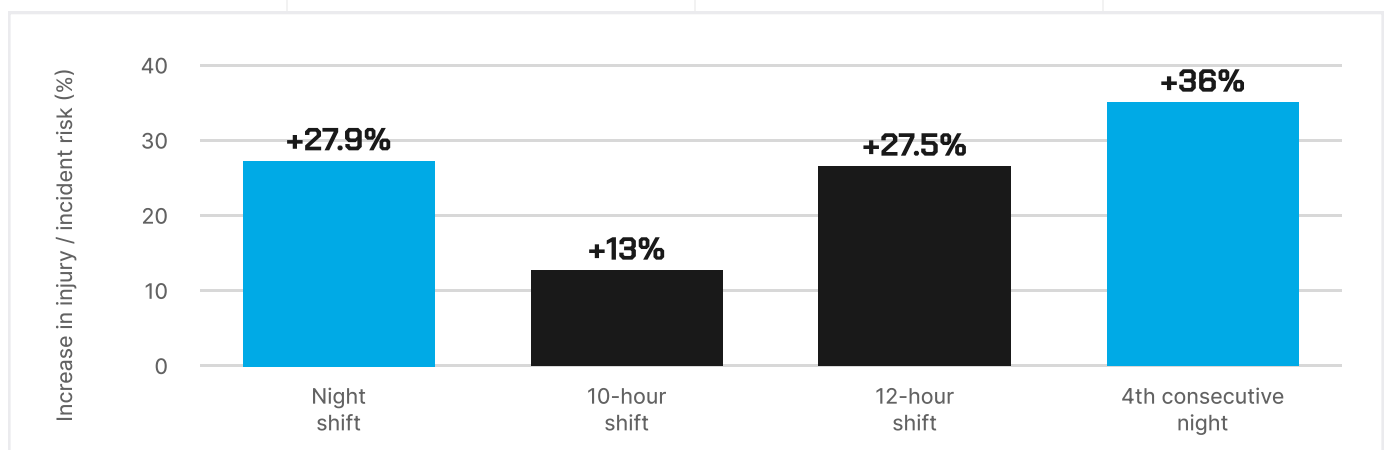


Figure 6. Increase in injury/incident risk by shift characteristic, relative to a morning/8-hour baseline. Source: Folkard & Lombardi, American Journal of Industrial Medicine 49(11), 2006.

The implication for a 24/7 Global Security Operations Center (GSOC) is direct. The overnight hours, when intrusion and insider risk are often highest, are precisely when the human operator is least cognitively capable, and the standard remedies (longer shifts to reduce handovers, consecutive nights to stabilize rosters) tend to make the impairment worse, not better.

06 · THE LABOR EQUATION

A high-turnover, hard-to-staff workforce is being asked to perform a task that even a stable, expert workforce could not perform at scale.

The cognitive constraints would be challenging even with an abundant, stable, well-rested workforce. The security-officer labor market is none of those things.

Scale, wages, and churn

The U.S. Bureau of Labor Statistics (BLS) counts roughly 1.26 million security guards (a figure for the narrow occupation, reported separately from the broader “Security Guards and Gambling Surveillance Officers” grouping). The median wage is about \$18.46 per hour. BLS projects little or no employment change through 2034, yet about 162,300 openings every year. Most, in the BLS’s own words, replace workers who transfer to other occupations or exit the labor force. That replacement churn is the quiet story of the industry. A widely cited figure puts annual turnover in contract security between 100% and 300%; it should be attributed carefully (it originates as a Service Employees International Union estimate without published methodology, not a government measurement), but the BLS openings-to-headcount ratio independently corroborates an exceptionally high rate of churn. Industry surveys reflect the strain: in one 2025 benchmark of physical-security operations, more than 40% of providers named turnover their single greatest challenge.

The arithmetic of 24/7 coverage

Continuous coverage is expensive before a single person is hired. Staffing one post 24 hours a day, 7 days a week requires 168 hours of coverage weekly against a 40-hour work week: a minimum of 4.2 FTEs per post, and closer to 5 once vacation, sick time, training, and turnover relief are included. (This is standard coverage arithmetic, not a claim attributable to any single study.) A facility that wishes to staff even a handful of monitoring positions around the clock is therefore committing to dozens of full-time hires, recruited and retained in the tightest part of the labor market, to perform a task the previous sections show degrades with exactly the fatigue this staffing model produces.

The conclusion the numbers force is that human live-monitoring does not scale economically any better than it scales cognitively.



07 · WHEN VIDEO FAILS TO BECOME OPERATIONAL INTELLIGENCE

Cameras routinely capture events that no human sees in time to act. The public record shows this is rarely a simple “operator wasn’t watching” failure. It is a breakdown somewhere along the chain from capture to response.

The most persuasive evidence that live monitoring fails at scale is that it has failed, repeatedly and on the record, in exactly the settings this white paper addresses. The cases below are drawn from official after-action reports, government audits, and major investigations, not vendor marketing. Each is classified by its dominant failure mode, because the lesson is not “hire better operators” but “the path from a recorded frame to an operational response has many points where it breaks.”

Washington Navy Yard, 2013: control-room communication and response-awareness breakdown

During the September 2013 mass shooting, a contract security operator sat in a control room with live feeds from 160 cameras that, in the words of the official record, were documenting the gunman’s movements in real time. The operator locked the control-room door, took cover, and did not relay what the cameras showed; responding officers were not even aware the camera control room existed and could not track the shooter in real time. The cameras worked perfectly. What failed was the human-to-responder communication and command layer: situational awareness was captured but never reached the people who needed it.

Metcalf substation, 2013: passive recording and delayed review

In April 2013, attackers fired on the Metcalf transmission substation operated by Pacific Gas and Electric (PG&E), damaging transformers and triggering an oil leak and overheating. The site’s cameras recorded the attack, but the footage was reviewed only afterward; the initial alerting came from equipment alarms and a neighbor’s 911 call, not from anyone watching the feeds. No arrests were made. The attack was consequential enough that federal regulators subsequently ordered a mandatory grid physical-security standard (the North American Electric Reliability Corporation Critical Infrastructure Protection standard, NERC CIP-014). This is the textbook signature of forensic-only video: a real-time event captured but processed too late to matter. (Documented by the Congressional Research Service (CRS), with operational detail from the Federal Energy Regulatory Commission (FERC), PG&E, and contemporaneous reporting.)

New York City Transit, 2024: camera-health, alarm-validation, and maintenance-governance failure

A 2024 New York State Comptroller audit of subway security examined a sample of 80 tunnel-intrusion alarms drawn from 164,649 such alarms logged in just six months. For 15 of the 80 sampled events, no live video was available or the cameras were not working; the remaining 65 did have working video. The audit also documented a CCTV repair that ran close to a year against a three-day standard. The point is not that “the cameras never work” (most did) but that even a minority of blind or unvalidated alarms, multiplied across 164,649 activations, represents a large population of intrusion signals reaching no reliable set of eyes, and that maintenance and validation governance had measurably slipped.

These are not stories about bad operators. They are stories about a model that places the entire burden of real-time detection and response on a human chain with many failure points: passive recording, communication breakdown, camera health, alarm validation, maintenance governance, and response workflow. Cameras may capture events, but without a reliable detection-and-response layer, video remains forensic rather than preventive.

08 · WHAT THE EVIDENCE SAYS ACTUALLY WORKS

The academic record draws a clear line: actively monitored video produces measurable prevention; passively recorded video largely does not.

It would be easy to read the preceding sections as an argument against surveillance itself. It is not. The research is in fact specific about when video creates security value and when it does not, and the dividing line is precisely the variable this white paper has been examining: whether anything is actually evaluating the imagery in real time.

The most authoritative synthesis is a 40-year systematic review with meta-analysis by Piza and colleagues, published in *Criminology & Public Policy* in 2019. Pooling decades of evaluations, it found CCTV associated with a modest overall crime reduction of roughly 13% (with the largest effects concentrated in car parks). Most importantly, it distinguished by monitoring mode: actively monitored CCTV produced a statistically significant crime reduction, while passively monitored systems generated non-significant, essentially null effects. An earlier UK Home Office evaluation of 13 systems reached a compatible conclusion (only two showed a statistically significant crime reduction, and in one of those the change was attributable to confounding factors), and framed recorded video's primary value as post-incident investigation rather than prevention.

This is the crux of the executive decision. The evidence says active monitoring works and passive recording largely does not. The same evidence says humans cannot sustain active monitoring at fleet scale. An organization that accepts both findings is led, by the data, to a single conclusion: the continuous-evaluation layer must be provided by something that does not fatigue, does not saturate at sixteen feeds, and does not learn to ignore alarms, with humans repositioned to do what only humans can.



09 · THE COST OF THE HUMAN CEILING

The title promises “what they cost.” Here is a transparent, conservative model of the labor and risk economics, built from clearly labeled assumptions, not vendor math.

The findings consolidate into a stack of costs, each individually documented and each compounding the others: fatigue worsens the vigilance decrement; alarm fatigue worsens both; turnover lowers the experience that might partially offset them. To make this concrete, the model below estimates the labor economics of attempting live monitoring at three enterprise scales. Every figure is illustrative and assumption-driven; the assumptions are stated so a reader can substitute their own.

COST LAYER	WHAT DRIVES IT	EVIDENCE ANCHOR
Missed events	Vigilance decrement, feed multiplexing, and inattentional blindness make real-time detection incomplete by design, especially at scale and overnight.	Mackworth 1948; Tickner & Poulton 1973; Simons & Chabris 1999
Wasted response	False-alarm rates of 94–98% consume officer and operator time and money, and breed the desensitization that causes real alarms to be missed.	Sampson / DOJ COPS 2007; Joint Commission 2013
Labor & turnover	~4.2 FTE per 24/7 post, drawn from a workforce with very high churn and modest wages, each departure lowering average experience and detection quality.	BLS 2024; IAPSC / Piza 2025; ASIS 2025
Fatigue risk	Night and long shifts impair operators to alcohol-equivalent levels and raise incident risk up to ~36%, concentrated in the highest-risk hours.	Dawson & Reid 1997; Folkard & Lombardi 2006

Shared assumptions (conservative)

ASSUMPTION	VALUE USED	BASIS / RANGE
Coverage	24/7/365 (8,760 hrs/yr per staffed seat)	Standard
FTE per continuously staffed seat	4.2 (≈5 with relief)	168 hrs/wk ÷ 40-hr week
Fully loaded annual cost per monitoring FTE	\$65,000	Base wage ~\$38k + ~70% load (benefits, supervision, training, recruiting); range \$55k–\$80k
Alarms reaching operators	~0.3 per camera per 12-hr shift	After basic filtering; conservative
Non-actionable share of alarms	95%	Analogy-informed: policing 94–98%, healthcare 85–99%
Average alarm triage time	3 minutes	Conservative mid-range
Labor removed by validated-event routing	80% of non-actionable triage	Illustrative target, not a guarantee

Three scenarios

Seats staffed reflects realistic GSOC seat counts, not full effective coverage, which is the point: organizations under-staff relative to the science, and the resulting cameras-per-operator ratio runs far above the ~16-feed effective ceiling.

	250 CAMERAS	1,000 CAMERAS	5,000 CAMERAS
Monitoring seats staffed (assumed)	1	2	4
FTE required (4.2/seat)	~4.2	~8.4	~16.8
Annual coverage labor (loaded)	~\$275K	~\$545K	~\$1.09M
Cameras per operator on shift	~250	~500	~1,250
vs. science-based effective ceiling	~16	~16	~16
Alarms per 12-hr shift	~75	~300	~1,500
Non-actionable alarms per day	~140	~570	~2,850
Wasted triage labor (loaded \$/yr)	~\$80K	~\$325K	~\$1.6M
Potential savings from validated routing	~\$65K/yr	~\$260K/yr	~\$1.3M/yr

What the model is really showing

Live-monitoring effectively, at the ~16-feed ceiling the science supports, would require roughly 16 simultaneous operators per 250 cameras (about 66 FTEs, well over \$4M/year) and is economically infeasible at fleet scale. So organizations do the rational thing and staff a few seats, which means the actual cameras-per-operator ratio sits 15–80× above the effective ceiling. The result is a large, mostly unwatched estate plus a meaningful annual spend on triaging non-actionable alarms. A validated-event routing layer does not eliminate the security team; it redirects the wasted triage labor (on the order of tens of thousands of dollars per year at 250 cameras to well over a million at 5,000) toward verification and response, and shrinks the unwatched gap.

Beyond the line items

Two costs resist tidy quantification but belong in any executive case. **Overtime and turnover burden:** at 100%+ annual turnover, recruiting and retraining recur constantly, and thin rosters force overtime that adds a 10–20% premium to the labor above. **Detection and response latency:** the most consequential cost is the incident discovered only on review. Validated-event routing is best understood as compressing mean time to detect (MTTD) and mean time to respond (MTTR) from the post-incident timescale toward the real-time one, an improvement we describe directionally rather than with a single promised number, because the realized figure depends on the deployment.



10 · THE PATH FORWARD: AGENTIC PHYSICAL SECURITY

The answer is not to ask people to watch harder. It is to give the continuous-evaluation task to a system built for it, one that perceives, reasons, and acts, and reserve for people the judgment only they can provide.

Every constraint in this white paper points in the same direction. Human attention is finite, quick to fatigue, easily saturated, and trainable to discount false alarms. Yet real-time, active monitoring is what makes surveillance prevent rather than merely record. Reconciling those two facts does not require asking operators to try harder. It requires moving the continuous-observation task off the part of the system least able to perform it, and onto a layer built for it.

That layer is what now defines the category: Agentic Physical Security. It marks the shift from passive recording to purpose-built AI that perceives, reasons, assesses, and responds to real-world threats as they unfold, at machine speed and enterprise scale, with human operators in the loop rather than in the bottleneck.

What an agentic layer changes, and what it does not

A reasoning-based AI layer is well matched to the work that defeats human operators, for one reason: it does not share their constraints. Its attentiveness at hour eight does not decay the way a person's does at minute twenty, so it is not subject to the vigilance decrement. It does not saturate at sixteen feeds; one system can evaluate thousands of streams at once. It does not get bored, distracted, or desensitized, and the night shift does not impair it. In human-factors terms, an agentic layer removes the three failure modes this research identifies as intrinsic to human monitoring: temporal decline, multiplexing collapse, and alarm-driven desensitization.

This is not automation. Automated systems execute predefined rules and trigger on raw motion. An agentic system interprets complex environments: it connects signals across cameras, doors, and sensors over time, weighs location, behavior, and intent to judge true criticality, and adapts as conditions change. It runs as a continuous loop, and each cycle compounds on the last.

01 · SEE	02 · THINK	03 · ASSESS	04 · ACT
Perceive Accurate perception at scale across video, access, and sensor data. Every configured feed, all day and all night.	Reason Continuous reasoning that connects signals over time, past the point where human detection collapses.	Validate Evaluating location, behavior, and intent to determine true criticality. Validation before escalation, easing the alarm-fatigue cycle.	Respond Initiating investigation, escalation, or policy-driven response. Surfacing only what matters, to the people who must decide.



That is a real and material advantage. It is not the same as perfection. The accurate statement is not that AI watches everything, but that an agentic layer continuously evaluates configured video streams within its operating envelope. Its performance still depends on camera quality, scene conditions, model quality, configuration, integration, and operational tuning. It produces false positives and false negatives. Poor lighting, occlusion, blind spots, and bad camera placement can defeat it; it can drift as environments change; and it depends on network health and on the workflows and governance built around it.

What AI does not remove

An honest case for agentic security operations names its limits up front. An agentic layer does not remove:

- **False positives and false negatives.** Detection is probabilistic; tuning trades one error type against the other.
- **Physical and optical limits.** Lighting, occlusion, blind spots, camera placement, and resolution still bound what any system can see.
- **Scene change and model drift.** Environments evolve; models must be monitored and retuned, or accuracy decays.
- **Infrastructure dependence.** Network outages, camera-health failures, and integration gaps degrade performance regardless of model quality.
- **The need for human judgment.** Context, intent, proportionality, escalation authority, and accountability remain human responsibilities.
- **Governance work.** Configuration, performance review, and oversight are ongoing obligations, not one-time setup.

Agentic AI changes which problems the security team spends its attention on. It does not make the team unnecessary, and it does not make the system infallible.

Human in the loop, not in the bottleneck

The point is not to remove people. It is to invert the workflow and divide the labor along the line of comparative advantage. Instead of asking operators to find the rare signal in an ocean of benign video, the agentic layer performs continuous evaluation and surfaces only validated, contextualized events for human judgment. The operator stays in the loop, but is no longer the bottleneck. Officers move from staring at walls of monitors to verifying, deciding, escalating, and responding: the parts of the job that genuinely require human reasoning and accountability, and where experience compounds rather than erodes.

DIVISION OF LABOR

The agentic layer perceives, reasons, assesses, validates, and, where policy allows, initiates response. Human operators verify, decide, command, and act on the events that genuinely require judgment.

How Ambient.ai applies this

Ambient.ai created the Agentic Physical Security category and delivers it through the Ambient Platform, the Reasoning AI Platform for Agentic Physical Security. The platform is a unified intelligence layer that works with an organization's existing cameras, sensors, and access control systems, turning passive infrastructure into an agentic security operation that perceives, reasons, and responds in real time. It does not rip out and replace what is already deployed. It makes that infrastructure smart.

At the core of the platform is Ambient Intelligence, the AI stack behind Ambient.ai's approach, and its reasoning core, Ambient Pulsar. Ambient Pulsar is the first always-on, edge-optimized reasoning Vision-Language Model (VLM) purpose-built for physical security, trained on more than one million hours of real-world enterprise video and run continuously on the Ambient Edge Appliance. It supplies the always-on perception and continuous reasoning the rest of the loop depends on.

The design maps directly onto the evidence in this white paper. It supplies the continuous, active-evaluation layer the research shows is effective. It is not subject to the vigilance decrement, to multiplexing collapse at sixteen feeds, or to alarm desensitization. And it repositions skilled officers onto verification, command, and response, where human judgment is decisive.

Organizations do not have to adopt all of this at once. Ambient.ai describes a staged Path to Agentic Physical Security, and teams can enter at the point that matches their priorities and maturity, from Agentic Monitoring through Agentic Response, expanding on a shared reasoning foundation over time.

The strategic case is deliberately narrow. It is not that AI is a better watcher than a person for twenty minutes, and it is not that AI prevents every incident. It is that an agentic layer removes the specific, well-documented human failure modes from the continuous-evaluation task, and in doing so makes the human security team more effective, not less.



11 · DEPLOYING RESPONSIBLY: PRIVACY, GOVERNANCE, AND TRUST

Moving humans up the stack and AI into continuous evaluation raises legitimate questions about privacy and oversight. A credible program answers them by design.

Recommending an AI-enabled evaluation layer carries an obligation to deploy it responsibly. The practices below are not legal boilerplate; they are the operating disciplines that keep an agentic security program trustworthy, auditable, and aligned with the organization's values and obligations.

The guidance in this section is deliberately vendor-neutral. It describes the practices any organization should expect from any Agentic Physical Security program, whatever platform delivers it, and it is written to be applied as evaluation criteria rather than as a description of one product.

- **Privacy by design.** Scope evaluation to security-relevant signals and configured streams; minimize collection; avoid repurposing security video for unrelated surveillance.
- **Human-in-command decision-making.** AI detects, reasons, and routes; consequential decisions (escalation, intervention, reporting) remain with accountable humans.
- **Auditability.** Maintain a reviewable record of detections, validations, routing decisions, and human actions, so any outcome can be reconstructed and examined.
- **Defined escalation policy.** Document who is notified, in what order, and on what criteria, so response is consistent and reviewable rather than improvised.
- **Data-retention discipline.** Set and enforce retention periods appropriate to the use and the law; do not retain footage indefinitely by default.
- **Role-based access.** Restrict who can view feeds, events, and recordings to those with a defined need, logged and periodically reviewed.
- **Performance review for false positives and false negatives.** Treat model accuracy as a governed metric, reviewed on a schedule, with tuning decisions documented.
- **Restraint on biometric identification.** Avoid facial recognition or other biometric identification unless it is explicitly authorized by applicable law, organizational policy, and customer requirement.
- **Clear stakeholder communication.** Where appropriate, communicate with employees, tenants, and the public about what the system does and does not do.

Done well, these practices are not a brake on the program. They are what makes it defensible to a board, a regulator, a works council, and the people who live and work inside the cameras' field of view.



12 · HOW TO KNOW YOU HAVE REACHED THE HUMAN CEILING

A one-page diagnostic for security leaders. If several of these indicators describe your operation, your live-monitoring model has likely passed the point where adding cameras or staff meaningfully improves real-time detection.

The “human ceiling” is not an abstraction; it shows up in measurable operational metrics. Use the following as a discovery instrument, for your own operation, or in conversation with one you are trying to understand. None of these is damning on its own; a cluster of them is the signature of a model that has outgrown its watchers.

DIAGNOSTIC METRIC	WHAT A “CEILING” READING LOOKS LIKE
Cameras per operator	Ratio far above ~16 effective feeds (commonly hundreds to thousands per seat)
Alarms per shift	Volume no operator could meaningfully triage in the shift
% of alarms requiring action	Very low (single digits), indicating a low-trust alarm stream
False escalation rate	High share of escalations that turn out to be non-events
Missed-event discovery rate	Important footage routinely found only during post-incident review
Average time to detection (MTTD)	Measured in hours or “on review,” not minutes
Average time to response (MTTR)	Lags detection because detection itself is late
Average alarm handling time	Long enough that a queue forms during busy periods
Camera uptime	Meaningful share of cameras offline or degraded at any time
Camera blind-spot rate	Known coverage gaps that are tolerated rather than closed
Overnight staffing burden	Thinnest coverage during the highest-risk hours
Operator turnover	High enough that experience never accumulates
Overtime dependency	Routine reliance on overtime to fill the schedule
% of video used only after incidents	Most of the estate functions as a forensic archive, not a preventive layer

If a leadership team cannot readily answer these, or answers them in the “ceiling” column, the live-monitoring model is the constraint, and no amount of additional capture changes that.

13 · WHERE THE HUMAN CEILING BITES HARDEST: BUYER ENVIRONMENTS

The same structural problem expresses itself differently across industries. This section maps the thesis to the environments where it matters most.

The human ceiling is universal, but its consequences and its highest-value fixes differ by environment. The table below is a discovery aid: it pairs each environment's reason live monitoring fails at scale with the event that matters most, what validated-event routing improves, and the executive pain point it addresses.

ENVIRONMENT	WHY LIVE MONITORING FAILS AT SCALE	EVENT THAT MATTERS MOST	WHAT VALIDATED ROUTING IMPROVES	PRIMARY EXECUTIVE PAIN POINT
Corporate campuses	Sprawling multi-site estates, thin after-hours staffing	Tailgating, after-hours intrusion, workplace-violence precursors	Continuous evaluation across all sites; only validated events reach the GSOC	Labor cost, workplace safety, response time
Data centers	Hundreds of cameras, near-zero tolerance for unauthorized access	Perimeter breach, unescorted access, tailgating into halls	Real-time validation of access anomalies; audit trail	Uptime, compliance, risk
Critical infrastructure	Remote, unmanned, geographically dispersed sites	Perimeter attack, sabotage, copper/equipment theft	Real-time detection at unmanned sites instead of after-the-fact review	Risk, regulatory compliance (e.g., NERC CIP), uptime
Healthcare	Open campuses, high foot traffic, alarm-saturated environment	Infant abduction, emergency-department violence, controlled-substance diversion, wandering	Cuts alarm noise; surfaces validated safety events	Workplace safety, compliance, labor
Higher education	Large open campuses, transient population, limited staff	After-hours intrusion, assault, property crime, protest safety	Continuous coverage of a sprawling estate without proportional headcount	Labor cost, safety, response time
Manufacturing & logistics	Large footprints, loading docks, high asset throughput	Theft, diversion, safety violations, unauthorized access	Validated detection across docks and yards; safety-event surfacing	Cost control, workplace safety, loss prevention
Transit & public sector	Vast estates (stations, tunnels, vehicles), public exposure, camera-health drift	Intrusion, trespass, platform incidents, infrastructure attack	Validates intrusion alarms and flags camera-health gaps in real time	Risk, public safety, uptime, compliance

14 · PUTTING THE THESIS TO WORK: DISCOVERY QUESTIONS

For security leaders assessing their own operations, and for teams helping them, these questions convert the white paper's argument into a practical conversation.

The strongest response to “we already have cameras” or “we already have guards watching them” is not a counter-assertion; it is a set of questions whose answers reveal the human ceiling in the operation's own numbers:

- How many of your cameras are actively monitored in real time, versus recorded for later review?
- How many alerts does your team receive per shift, and what percentage lead to action?
- How often do you discover the important footage only after the incident?
- How do you know which cameras are not being watched, or are not working, right now?
- What are your current mean time to detect and mean time to respond?
- What would change if every camera could be continuously evaluated for validated risk signals, and your officers only saw the events that mattered?

These questions are not rhetorical. In most large operations the honest answers describe exactly the structural mismatch this white paper documents, and they point, on their own, toward a machine-supported detection layer with humans elevated to judgment and response.

15 · CONCLUSION

The constraint was never the technology of capture. It was always the human ceiling on attention.

Enterprise security has spent two decades solving the capture problem (more cameras, higher resolution, longer retention) while leaving the detection problem to a human attentional system that has not changed since Mackworth measured its limits in 1948. The result is the structural mismatch documented throughout this white paper: vast estates of live imagery, a workforce that fatigues and turns over, an alarm environment that trains people to stop listening, and a public record of consequential events that cameras captured but that no one acted on in time.

None of the underlying science is controversial. The vigilance decrement, the working-memory ceiling, inattention blindness, alarm fatigue, and fatigue-as-impairment are among the best-established findings in their respective fields. What this white paper contributes is their assembly into a single, executive-level picture of a problem the security industry has tended to treat as a matter of operator diligence rather than human biology, and the distinction matters, because the two diagnoses lead to opposite remedies.

The path forward follows from the evidence rather than from any vendor's preference. Continuous, active evaluation is what makes surveillance prevent harm; humans cannot provide it at fleet scale; therefore the continuous-attention layer should be machine-provided, within its operating envelope and under human command, with people elevated to the judgment, escalation, and response work that is irreducibly theirs. This is the essence of Agentic Physical Security. Organizations that make this shift do not weaken their security teams. They align the work with the capabilities of the people doing it, and close the gap between what their cameras can see and what their operations can act on.

REFERENCES

All sources below were verified against primary documents or originating institutions; access date 26 June 2026.

- Green, M. W. (1999). The Appropriate and Effective Use of Security Technologies in U.S. Schools: A Guide for Schools and Law Enforcement Agencies. Sandia National Laboratories for the National Institute of Justice, U.S. DOJ. NCJ 178265. nij.ojp.gov
- U.S. Government Accountability Office (2002). National Preparedness: Technologies to Secure Federal Buildings. GAO-02-687T. gao.gov/assets/a109302.html
- Mackworth, N. H. (1948). The breakdown of vigilance during prolonged visual search. *Quarterly Journal of Experimental Psychology*, 1(1), 6–21. DOI: 10.1080/17470214808416738.
- Klein, R. M., & Feltmate, B. B. T. (2025). The vigilance decrement: its first 75 years. *Frontiers in Cognition*, 4, 1632885. DOI: 10.3389/fcogn.2025.1632885.
- Tickner, A. H., & Poulton, E. C. (1973). Monitoring up to 16 synthetic television pictures showing a great deal of movement. *Ergonomics*, 16(4), 381–401. DOI: 10.1080/00140137308924529.
- Donald, F. M., & Donald, C. H. M. (2015). Work exposure and vigilance decrements in closed circuit television surveillance. *Applied Ergonomics*, 47, 220–228. DOI: 10.1016/j.apergo.2014.10.001.
- Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: sustained inattention blindness for dynamic events. *Perception*, 28(9), 1059–1074. DOI: 10.1068/p281059.
- Drew, T., Vö, M. L.-H., & Wolfe, J. M. (2013). The invisible gorilla strikes again: sustained inattention blindness in expert observers. *Psychological Science*, 24(9), 1848–1853. DOI: 10.1177/0956797613479386.
- The Joint Commission (2013). Sentinel Event Alert, Issue 50: Medical device alarm safety in hospitals. April 8, 2013. jointcommission.org
- ECRI Institute. Top 10 Health Technology Hazards (annual reports). Clinical-alarm hazards ranked #1 on the lists for 2012 through 2015.
- Drew, B. J., Harris, P., Zègre-Hemsey, J. K., et al. (2014). Insights into the problem of alarm fatigue with physiologic monitor devices: a comprehensive observational study of consecutive ICU patients. *PLOS ONE*, 9(10), e110274. DOI: 10.1371/journal.pone.0110274.
- Sampson, R. (2007). False Burglar Alarms, 2nd ed. Problem-Oriented Guides for Police, No. 5. U.S. DOJ Office of Community Oriented Policing Services / ASU Center for Problem-Oriented Policing. popcenter.asu.edu
- U.S. Bureau of Labor Statistics (2025). Occupational Outlook Handbook: Security Guards (May 2024 OEWS, SOC 33-9032; 2024–34 projections). bls.gov/ooh/protective-service/security-guards.htm

- Dawson, D., & Reid, K. (1997). Fatigue, alcohol and performance impairment. *Nature*, 388(6639), 235. DOI: 10.1038/40775.
- Williamson, A. M., & Feyer, A.-M. (2000). Moderate sleep deprivation produces impairments in cognitive and motor performance equivalent to legally prescribed levels of alcohol intoxication. *Occupational and Environmental Medicine*, 57(10), 649–655. DOI: 10.1136/oem.57.10.649.
- IHS Markit / Omdia, reported by CNBC (2019). One billion surveillance cameras will be watching globally in 2021. *cnn.com*. Installed base reached ~770M (2019) and surpassed one billion (2021); the 2027 figure is a projection.
- Piza, E. L. (2025). Video Surveillance (Evidence-Based Standard of Practice EBSP-25-01). International Association of Professional Security Consultants. *iapsc.org*
- Al-Shargie, F., et al. (2019). Vigilance decrement and enhancement techniques: a review. *Brain Sciences*, 9(8), 178. PMC6721323.
- Miller, G. A. (1956). The magical number seven, plus or minus two. *Psychological Review*, 63(2), 81–97.
- Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24(1), 87–114. DOI: 10.1017/S0140525X01003922.
- Folkard, S., & Lombardi, D. A. (2006). Modeling the impact of the components of long work hours on injuries and “accidents.” *American Journal of Industrial Medicine*, 49(11), 953–963. DOI: 10.1002/ajim.20307.
- Briscoe, S. (2025). Guarding companies face the challenge of high turnover. *Security Management (ASIS)*, reporting the Trackforce 2025 Physical Security Operations Benchmark Report; SEIU 100–300% turnover estimate via Bitzer, E. (2006). *asisonline.org*
- Piza, E. L., Welsh, B. C., Farrington, D. P., & Thomas, A. L. (2019). CCTV surveillance for crime prevention: a 40-year systematic review with meta-analysis. *Criminology & Public Policy*, 18(1), 135–159. DOI: 10.1111/1745-9133.12419.
- Gill, M., & Spriggs, A. (2005). Assessing the impact of CCTV. Home Office Research Study 292.
- Metropolitan Police Department (2014). After Action Report: Washington Navy Yard, September 16, 2013. July 11, 2014. *mpdc.dc.gov*
- Parfomak, P. W. (2014). Physical Security of the U.S. Power Grid: High-Voltage Transformer Substations. Congressional Research Service, R43604. (Attack-and-CIP-014 narrative; granular operational specifics trace to FERC/PG&E/press reporting.)
- New York State Office of the State Comptroller (2024). MTA–New York City Transit: Selected Aspects of Safety and Security in Subway Facilities. Report 2022–S-20, August 2024. *osc.ny.gov*

Prepared by Jody Russell, AI Evangelist and Solutions Engineer, Ambient.ai · June 2026. This document is provided for informational purposes. Statistical claims are attributed to the primary sources listed above; readers are encouraged to consult them directly.

