From Weak Signals to Decisions: Simulation-Based Training as a Response to the Early-Warning Dilemma in Crisis Management

A Case Study of Exercise ILVES 2026 (Estonia)

Aigo Allmäe

Estonian Academy of Security Sciences, Tallinn, Estonia
aigo.allmae@sisekaitse.ee

Abstract. Crisis management guidance tells decision-makers to act on weak signals – small, unclear hints that something may be going wrong. In practice this is very hard to do. Acting early means spending money and political capital on evidence that is not yet convincing, while waiting means acting once the crisis has started and the cheap options have gone. This paper calls this problem the early-warning dilemma. It argues that the problem is usually treated as a shortage of information when it is really a shortage of practice. Drawing on work about weak signals, early warning and cognitive bias, it sets out four pressures that push decision-makers towards acting too late. The paper then examines ILVES 2026, the largest crisis exercise ever held in Estonia (8–12 June 2026, more than 150 organisations, about 4,000 participants), described here by one of its organisers. The exercise was built on a six-step chain that begins with noticing change. Its scenario deliberately withheld a complete picture, organisations were assessed on paper before the week began, and assessors asked participants why they had decided as they did and checked the answers. The control team worked to a rule of keeping enough pressure on players that they did not wait for perfect information. The case shows a national exercise designed around the dilemma rather than around a script. Two things stay out of reach in a live exercise of any size, and immersive training is well placed to cover them. Five design principles with measurable indicators are proposed.

Keywords: Weak Signals, Early Warning, Crisis Management, Decision-Making Under Uncertainty, Simulation-Based Training, Immersive Learning, Comprehensive National Defence.

1 Introduction

Everyone agrees that noticing a crisis early is better than noticing it late. The earlier a problem is seen, the more options are still open and the less they cost. This is why foresight work puts so much weight on scanning for emerging issues and weak signals [1, 2], and why early-warning systems exist for floods, epidemics, armed conflict and financial instability [3].

© The Author(s) 2026
K. Hollá et al. (eds.), *Proceedings of the SAFAR Scientific Conference 2026 – Building Awareness Through Immersive Learning (SSC-BAIL 2026)*, Atlantis Highlights in Social Sciences, Education and Humanities 57,
https://doi.org/10.2991/978-94-6239-817-7_12

For thirty years researchers have described a warning–response gap. The warning is produced, it reaches the right desk, and nothing follows [4]. Two explanations dominate: one says the forecasts are not accurate enough, the other that the warning does not survive the journey through the bureaucracy. Both are true, and both have led to real improvements, but neither says much about what happens inside the decision itself.

This paper argues that the main obstacle sits in the decision, and that it takes the form of a dilemma. A weak signal is incomplete, unclear and buried in noise [5]. What makes it weak, above all, is that it can be read in more than one way [5]. To act on such a signal, or on a number of them, someone has to spend money, disrupt normal work and accept political risk on evidence that does not justify any of it. To wait until the evidence does justify it is to wait until the signal is no longer weak, which is to say until the crisis has already begun. A person who behaves sensibly at every step will therefore act too late, and a person who acts early will usually look, afterwards, as though they overreacted.

If that is right, the problem will not be fixed by better sensors or shorter reporting lines; it has to be lived with. People and teams need practice at acting on thin evidence, at keeping more than one explanation alive at once, and at carrying the blame that follows a false alarm. Skills come from practice with feedback, which is what training is for.

The paper asks three questions. First, what exactly is the early-warning dilemma, and where in the warning chain does it bite? Second, how far can a large national crisis exercise train the skills it demands, and what stays out of reach? Third, what should immersive training environments look like if they are to cover the rest? Sections 2 and 3 set out the argument. Section 4 reviews what is known about simulation-based and immersive learning. Section 5 explains the method, Sect. 6 presents the ILVES 2026 case, Sect. 7 proposes five design principles, and Sect. 8 concludes.

2 Weak Signals and the Early-Warning Chain

2.1 What a Weak Signal Is

Ansoff brought weak signals into strategic management as an alternative to planning that waits for strong ones. His idea was a graduated response, sized to the strength of the information actually in hand [6]. Lesca and Lesca describe a weak signal as a piece of data that looks unimportant, or that is lost in a mass of raw data; it becomes information only when someone happens to notice it [5]. What marks a weak signal, on a related reading, is that it supports competing interpretations [5]. Ilmola and Kuusi stress that the organisational meaning of such information is very hard to pin down early on [7]. Kuosa notes that its value is assigned by the observer, and usually appears only once several signals are grouped together [8]. Schoemaker and colleagues describe fragments that look random until a different frame, or a link to another fragment, turns them into part of a pattern [9]. Hiltunen separates the signal, the issue and the interpretation, which explains how the same signal can be strong in one respect and weak in another [10].

For training design, the definitional debate matters less than the working properties on which these authors agree. Table 1 lists those properties.

Table 1. What makes a signal weak and what a crisis system must be able to do about it.

What makes it weak	What that means for the observer	What the system must be able to do
It is incomplete	You only ever have part of the picture	Let people report before they are sure
It is buried in noise	Most people will not see it at all	Hold attention when data volume is high
It can be read in several ways	People will disagree about what it means	Keep competing explanations alive
It is unfamiliar	Nobody is looking for it, so nobody finds it	Look sideways, not only where told to look
It looks pointless	What strikes one person means nothing to the next	Spread judgement widely; keep the reporting threshold low
It looks unreliable	The source may not be trustworthy	Weigh weak evidence instead of discarding it
It stands alone	There is no file to put it in	Bring fragments together across fields
It is irregular	It has no frequency worth measuring	Do not dismiss it for being a one-off

Not one of these properties is addressed by a piece of technology alone; each is handled by trained people working inside a routine that expects this kind of information.

2.2 From Signal to Warning: Where the Chain Breaks

A single weak signal is rarely enough to act on. Lesca and Lesca describe turning one into an early-warning signal as amplification: a process of interpretation and shared sense-making that uses reasoning, rules of thumb, discussion among several people, a search for further information and a certain amount of imagination [11]. Their Puzzle method turns this into a facilitated group session in which fragments collected by scattered observers are laid out and fitted together until a plausible, supported hypothesis appears. The wider VASIC process places that session inside a cycle of setting the scanning target, individual tracking, passing information upward, circulating it and interpreting it together [11].

Two things about this account matter here. First, amplification is a team activity: the person who notices a fragment is usually not the person who can interpret it, and neither of them is the person allowed to act. Second, it breaks easily at both ends. Lesca and Lesca point to the problem of setting the scanning target – too narrow and the system starves, too broad and the useful fragments drown [5]. Horizon scanning runs into the same difficulty, which is part of why it is hard to do well and hard to defend to the people paying for it [12, 13].

Woocher splits early warning into three parts: estimating which risks are largest and when, analysing what the threat actually is and describing plausible scenarios, and passing that analysis to people who can act. The receiving side then has two jobs of its own: it has to judge the size and nature of the threat for itself, and it has to decide whether anything new should be done [4]. The natural-hazard literature describes a similar chain of monitoring, forecasting and communication, and adds a blunt condition: such a system works only where officials have the will and the standing resources to act, and training is part of the system rather than an extra [3].

The break between the two halves is well documented. George and Holl named the warning–response problem in preventive diplomacy [4]. Disaster risk reduction meets the same asymmetry in the question of whether a warning actually reaches, and is believed and understood by, the recipient [3]. The same discontinuity appears inside the individual, where situation-awareness research separates seeing something from understanding it and from working out where it is heading [14], and inside the group, where research on high-reliability organisations describes the habits of watching for small failures and resisting easy explanations [15].

3 Four Pressures That Delay the Decision

The warning–response gap is not mainly a transmission failure. It is what you see when four separate pressures all push a decision in the same direction.

The first is evidence pressure. Evidence about a developing problem gets stronger over time. The value of doing something about it falls over the same period. The window in which action is cheap and effective is exactly the window in which the evidence is thinnest. Ansoff's graduated response is a partial answer, but it assumes an organisation prepared to spend on a state of knowledge that does not yet justify spending [6].

The second is bias pressure. Woocher's review suggests that the two sides of the chain are pulled in opposite directions. Analysts, who work with what comes easily to mind and who spend their days writing detailed scenarios of the events they are paid to anticipate, tend to overestimate risk; describing a scenario in detail makes it feel more likely [4, 16]. Decision-makers, who see prevention as a certain loss now against an uncertain loss later, tend to wait, because people take risks to avoid losses [4, 17]. One side pushes the alarm up, the other pushes the response down, and the gap widens from both ends. Knowing about this does not help much on its own: Heuer's finding that awareness of a bias does not by itself improve perception is probably the single most important result for anyone designing exercises and training [16].

The third is blame pressure. Hindsight works unevenly: it is much stronger for things that happened than for things that did not [4, 18]. A crisis that occurred is judged, afterwards, to have been obvious. A crisis that was prevented leaves nothing to compare against, so nobody gets credit for preventing it. The result is a scoreboard on which a missed warning is punished later, a false alarm is punished at once, and success is invisible. Under those rules, the safe career move is to escalate late.

The fourth is structural pressure. Noticing happens at the edges of an organisation or a society; deciding happens at the centre. The fragments sit with people who cannot

interpret them, the interpreting is done by people who did not see them, and the decision belongs to people who have done neither. Every handover is a filter, and every filter is set to reject exactly the things listed in Table 1: the incomplete, the unclear and the apparently pointless.

Together these pressures mean that the dilemma cannot be removed. It can only be handled better or worse. That is what turns an analytical problem into a training problem.

4 Why Training, and Why Immersive Tools

Three arguments point towards training rather than towards better analytical technology.

The first is that the gap is one of experience, not of knowledge. Nobody in this field needs to be told that early action is valuable. What people lack is any real sense of what it feels like to commit on thin evidence, and of what follows when they do. Fischhoff's point that training supplies the feedback which working life does not is particularly relevant here, because genuine crises are far too rare for a single career to produce that feedback by itself [4, 18].

The second is that the skills involved are the non-technical ones: reading a situation, holding a hypothesis, saying out loud how confident you are, and coordinating with others. Kolb's experiential cycle and the cognitive-affective model of immersive learning both predict that such skills are learned by doing something and then reflecting on it in a structured way, not by being told about it [19, 20].

The third is that there is now evidence for the effect. In a survey of 352 emergency-response professionals, virtual-reality training predicted better crisis decision-making, and the sense of immersion accounted for roughly half of that effect [21]. Evain and colleagues describe the same shift in practice, away from traditional crisis exercises and towards immersive ones [22], and similar results have been reported for military decision training [23].

The same evidence comes with a warning. In that study the benefit of immersion was largest in shorter sessions and faded as sessions ran longer [21]. This fits work on cognitive load showing that long spells in immersive environments raise mental effort and reduce what is retained [24, 25]. More immersion is not automatically better. It is a resource, and it should be spent where nothing else will do the job.

For this problem, there is one thing simulated environments can do that neither a lecture nor a conventional tabletop exercise can. They let the designer set the amount of noise around a cue. In a tabletop exercise an inject arrives labelled as an inject, so its importance is guaranteed by the way it is delivered. In a simulated environment the same cue can be placed among competing, irrelevant and contradictory material, and how easy it is to spot can be varied on purpose. Without that, noticing cannot really be practised, observed or assessed.

5 Method

The study combines a narrative review of the literature with an exploratory single-case study. The review covers weak signals, early warning, foresight and cognitive bias; the four pressures in Sect. 3 are drawn from it, and it supplies part of the basis for the design principles in Sect. 7.

Case Selection. The case is the Estonian comprehensive national defence exercise ILVES 2026 [26, 27]. It was chosen for four reasons. It is recent, the main week having ended in June 2026. It is the largest exercise of its kind ever run in Estonia. It cuts across sectors, involving state agencies, municipalities, companies and civil society, and its scenario is a hybrid, non-military one, which means the ambiguity in it is not a trick of exercise design but a real feature of the threat being exercised. Finally, Estonia's comprehensive defence model, in which many organisations hold crisis duties and few hold decision rights, is a clear example of the structural pressure described in Sect. 3.

Data Sources. Two kinds of material were used. The first is public documentation: the Government Office pages on the exercise in Estonian and English, government press releases of 8 April and 12 June 2026, published accounts by participating municipalities, and the relevant legislation [26–32]. The second is the author's own knowledge of the exercise.

The Author's Role. The author was one of the organisers of ILVES 2026 and a member of the team that ran it, with access to the design logic, the scenario brief and the work of the exercise control centre. This position is what makes the case possible to describe, and it is also the study's main source of potential bias. Two steps were taken to limit that risk. First, statements about the exercise are kept to design decisions that can be set out plainly and checked by other members of the exercise staff, rather than to judgements about how well any organisation performed. Second, no claim is made about effectiveness: the paper does not argue that ILVES improved anyone's ability to read weak signals. It argues that the design addressed the warning chain in a particular way, that this is worth describing because such designs are rarely written down, and that certain parts of the chain stay out of reach in a live exercise of any size.

Analysis. Documents were coded against categories set in advance from the framework in Sects. 2 and 3, in the manner of directed content analysis [33], and the framework was revised where the material showed a category to be wrong.

Findings and Interpretation. A consistent distinction is kept between three kinds of statement. Descriptions of what the exercise was and how it was designed (Sects. 6.1 and 6.2) rest on public documents and on design decisions verifiable by other members of the exercise staff. The comparison between the design and the literature, and the

account of what stays out of reach (Sects. 6.3 and 6.4), are the author's interpretation of that material. The five principles in Sect. 7 are the author's proposals for future training design; they draw on the case but deliberately go beyond it, and Sect. 7 states for each principle where it comes from.

Limitations. Written control documentation, including scenario files, inject lists and evaluation reports, was not publicly available when this paper was written. No systematic observation protocol, interview series or participant survey was used, so the case is exploratory. A single case supports analytic rather than statistical generalisation [34]. The findings are therefore best read as hypotheses about a type of exercise, stated clearly enough to be tested.

6 The Case: Exercise ILVES 2026

6.1 Background and Conduct of the Exercise

Estonia organises crisis preparedness around comprehensive national defence: security is not the army's business alone, and state agencies, local governments, companies and communities each play a part in it [26]. The legal framework was rewritten during the period studied here. The Crisis Situation and National Defence Act (KORKS) was passed by the Estonian Parliament on 2 June 2026 and comes into force on 1 October 2026. It merges the former National Defence Act, State of Emergency Act and Emergency Act into one law and introduces two unifying terms, the crisis situation and the permanent crisis task [29]. Changes to the Emergency Act in October 2024 had already raised the number of vital services to twenty-two, supplied by more than four hundred companies, each of which must keep its service running in a crisis [26]. Estonian law requires a nationwide complex crisis-management exercise at least once every four years; the previous one, CREVEX, was held in 2023 [26].

A law that gives permanent crisis duties to several hundred organisations creates, without meaning to, a very wide network of potential observers. It also adds a great many organisational borders that a small, uncertain observation has to cross before it reaches anyone who can act on it.

ILVES 2026 was run by the Government Office together with the Ministry of the Interior. The main week ran from 8 to 12 June 2026, with sub-exercises and preparation spread across the year: thirty-eight sub-exercises and sixty-nine training events were planned [28, 31]. More than 150 organisations took part, among them the Office of the President, the Parliament of Estonia, ministries and agencies, local governments, vital service providers, churches and congregations, non-governmental organisations, the Defence Forces and the Defence League, together with private partners such as water utilities and retail chains. About 4,000 people took part, against an initial plan of 3,000 [27].

The scenario placed Estonia under hybrid attack by a hostile state trying to undermine constitutional order by non-military means. Participants dealt with public disorder, information operations, sabotage and sudden attacks meant to disrupt daily life,

spread fear and weaken trust in the state [27]. Several hundred linked incidents were run across the country during the main week. At least one municipal crisis committee worked through a water-supply failure and contamination while running on generators, with normal internet down and satellite phones and radios in use [32]. A separate cash-circulation exercise involved nineteen municipalities and the Ministry of Finance, a nationwide sheltering exercise was held on 10 June, and assigned assessors observed crisis communication throughout the week [30].

6.2 How the Exercise Was Designed

The published objectives were to practise command and coordination, assess capabilities, strengthen cooperation, raise readiness and improve societal resilience [26]. A list of that kind, however, says little about how the exercise actually worked. Four features of the design matter here.

A Chain That Begins with Noticing. Every participating organisation was expected to work through six steps: notice change; build situational awareness; anticipate what comes next; find measures that take back the initiative; make decisions, including creative ones; and resolve the situation. The first step was not to "respond to the event"; it was to "see that something has changed", which is where the dilemma in Sect. 3 begins. Once an organisation had received and interpreted an input, it worked through the same impact assessment: resources, people, time, management, communication, and the risks that would come next.

A Scenario Built to Withhold. The design team set out, in advance and in writing, what the scenario would and would not do for the players. Table 2 reproduces it. The right-hand column is the unusual part: the scenario was built to withhold things on purpose.

Table 2. What the ILVES 2026 scenario was designed to do, and what it was designed not to do.

What the scenario does	What the scenario does not do
Shows what hybrid attack actually looks like	Give a full description of the situation
Gives the whole week a shared background	Give the players the answers
Pushes organisations towards cooperation, information sharing and decisions	Remove contradictory information
Presents unclear signals that have to be interpreted	Solve tasks or make choices on the players' behalf
Makes players think about what might happen next	Make the crisis clean, linear or comfortable

A Control Rule About Pressure. The exercise control centre followed the answers players produced, and how things played out in the field, and adjusted the scenario as

it went, so that the situation developed in response to real choices. The organisers described their own job in a single sentence: keep enough pressure on the game that the players do not wait for perfect information, but move to a decision. Every activity was assessed, and feedback came immediately rather than months later.

Assessment That Asked Why. Assessors did not only record what organisations did. Working with the control centre, they asked participants why they had acted as they did and why they had believed what they believed, and validated those answers against what had, in fact, been injected. This is the part of the design that speaks most directly to Sect. 3: the dilemma is about judgement rather than procedure, and judgement can only be assessed by asking for it.

A Baseline Taken on Paper First. Before the main week, participating organisations went through a static assessment in which their crisis preparedness was analysed on paper: plans, roles, resources and arrangements as written down. The exercise then put the same organisations under pressure. Having both allows a rare comparison between what an organisation had declared it would do and what it actually does when the picture is unclear and time is short.

Read against Sect. 3, the control rule is the early-warning dilemma written down as an instruction to the control room. The organisers had identified the tendency to wait for certainty as a behaviour the training was intended to counter, and had made this a standing objective.

6.3 The Design Chain and the Literature

The six steps were not derived from the academic literature on weak signals; they were built by practitioners for a national exercise. It is therefore worth noting how closely they follow the chain described in Sect. 2. Table 3 sets the two side by side. The comparison is the author's interpretation rather than a finding of the exercise itself.

Table 3. The ILVES design chain and its counterpart in the literature.

Step in the ILVES chain	Counterpart in the literature
1. Notice change	Detection: the fragment that happens to catch someone's attention and so becomes information [5]
2. Build situational awareness	Perception and comprehension [14]; amplification of a weak signal through shared sense-making [11]
3. Anticipate developments	Projection [14]; describing plausible scenarios rather than predicting events [4]
4. Take measures that regain the initiative	Graduated response sized to the strength of the information in hand [6]
5. Take decisions, including creative ones	The response half of the warning chain: judge the threat, then decide what to do [4]
6. Resolve the situation	Outcome and, through the "next risks" question, a return to step 1

The match is close enough to be useful in both directions. For practitioners, it suggests that the six steps are not just a sensible sequence but a well-grounded one. For researchers, it is a reminder that exercise designers working on real problems tend to arrive at the same structure that the literature describes, usually without citing it, and that these designs are rarely written down where anyone else can learn from them.

6.4 What Stays Out of Reach

Two things remain difficult, and they are hard because of the form of a live national exercise rather than because of anything the organisers did. A third is not out of reach at all, but is a question about measurement.

First, the noise has a ceiling. The scenario was deliberately incomplete and contradictory, which is as far as a live exercise can go. What it cannot do is make participants uncertain that anything is happening at all. Everyone knows the week is the exercise; attention is already switched on, and it is pointed in roughly the right direction. In real life the hardest part of step one is that nobody has told you to start looking. That specific difficulty cannot be reproduced when the calendar says the exercise begins on Monday.

Second, the pressure runs one way. Keeping players from waiting for perfect information is the right rule for the failure this paper is about, and it was deliberate. It does mean, though, that only one of the two possible mistakes carries a cost. Raising the alarm without cause spends resources, damages credibility and can be exactly what an adversary conducting a hybrid campaign wants. If the exercise never prices that, participants practise moving early without practising the judgement of when not to.

Third, judgement was assessed but not yet measured. Asking an organisation why it decided as it did, and checking the answer against what was actually injected, is exactly what the literature on debiasing asks for, and it was done. Turning that into a measurement is a further step. It needs the reasoning to be captured in a standard form, so that interpretation can be compared between organisations and followed from one exercise cycle to the next, and it needs the same situation to be run more than once. A national exercise happens once every few years and never twice the same way. This is where a

live exercise and a repeatable immersive module help each other: the exercise supplies scale and realism, the module supplies the second run.

7 Five Design Principles for Immersive Training

ILVES already applies much of what follows, in a live setting and at a national scale. It withholds the answers, it spreads the problem across organisations that have to talk to each other, and it asks participants to account for their reasoning. What immersive tools add is not novelty but control and repetition: the ability to set the difficulty deliberately, run the same situation again with a different group, and measure what changed. Table 4 sets out five principles. Each answers one or more of the pressures in Sect. 3, and each has an indicator that can be measured.

The principles differ in how directly they derive from the case. P2, P3 and P5 generalise design features that ILVES 2026 already implemented in live form: the deliberately ambiguous scenario, the spread of duties across organisations, and assessment that asked why. P1 and P4 respond to the two limits identified in Sect. 6.4 – the noise ceiling and the one-way pressure – and are therefore the author's proposals for what a live exercise cannot do; they go beyond the case and remain to be tested.

Table 4. Design principles for immersive training on weak signals, with indicators.

Principle	What it requires	Answers to	Indicator
P1. Build in the noise	Hide cues among irrelevant, repeated and contradictory material, in a setting where the participant has not been told that anything is happening	Evidence pressure; noise ceiling	How many cues are found, and how quickly, at different noise levels; how many false alarms
P2. Keep cues open to more than one reading	Write every important cue so that at least two sensible readings fit, and withhold the answer until after the decision	Bias pressure	How many readings each cue produces; how often teams change their minds when new evidence arrives
P3. Spread the cues across roles	Give no single participant enough to work it out alone, and reward passing fragments upward	Structural pressure	Share of important cues that start at the edge; time from first sighting to the centre knowing
P4. Make both mistakes cost something	Show the price of a false alarm as well as the price of a missed warning, inside the scenario itself	Blame pressure; one-way pressure	Decisions scored for both kinds of error; credit for a well-argued decision not to escalate
P5. Debrief the thinking, not only the actions	Ask why people believed what they believed, and make their confidence explicit	All four pressures; comparison across cycles	Share of debriefing spent on reasoning; how well confidence matches accuracy across runs

Two constraints follow from Sect. 4. Because the benefit of immersion fades as sessions get longer [21, 24], these principles point to short, repeated, deliberately unclear episodes rather than long immersive scenarios. And because they are about judgement rather than procedure, the reflective part is not an add-on at the end; it is half the exercise. Foresight practice already supplies tools for it. The futures triangle, which separates the push of the present, the pull of the future and the weight of the past, and the sequence of futures questions that ends by naming hidden assumptions, both work well as debriefing structures for interpretation [1, 2]. The Puzzle method offers a matching structure for putting fragments together during the exercise itself [11].

There is a practical implication for how exercises are planned. The two forms are complements, not rivals. A national exercise can do what no laboratory can: put several thousand real people, in their real jobs, under real time pressure, across a whole country. What it cannot do is repeat itself, vary one condition at a time, or start before the participants know it has started. The preparation phase of a national exercise cycle – in ILVES, the sixty-nine training events and thirty-eight sub-exercises spread through the year [31] – is where small groups can be put repeatedly into noisy, unclear situations, with the noise level set on purpose and the results measured. The main week is then where organisations apply the same chain at full scale. ILVES already has the two ends of an evaluation design in place, in the paper assessment taken beforehand and the validated reasoning collected during the week; a repeatable module in between would complete it.

In plain terms, the five principles ask designers of large-scale crisis training to do five things: mix the important information with noise, so that finding it is part of the task; keep each clue open to more than one reasonable explanation; spread the information across different people, so that they have to share what they have and work it out together; show, inside the exercise, that acting too early and acting too late both have consequences; and, in the debrief, not stop at what people did but ask why they believed something was happening and why they decided as they did.

8 Conclusion

The dilemma of using weak signals is not a shortage of information. It is what happens when four pressures – evidence, bias, blame and structure – all push a decision the same way, towards acting later than would be sensible. No amount of engineering removes it, and the two usual remedies, better models and better warning channels, leave it untouched.

It can, however, be trained for, and ILVES 2026 shows a national exercise doing so on purpose. The design chain begins with noticing change rather than with responding to an event. The scenario was written to withhold a complete picture, to leave contradictions in place and to keep the crisis untidy. Preparedness was assessed on paper before the week, and during it assessors asked organisations why they had decided as they did and checked those answers against what had actually been injected. Most tellingly, the control room worked to a rule aimed squarely at the dilemma: keep enough pressure

on the players that they do not wait for perfect information. Designs of this kind are common in practice and rare in print, which is reason enough to describe this one.

Two things stay out of reach in a live exercise of any size: the noise has a ceiling, because everyone knows the exercise has begun, and the pressure runs one way, so only the cost of acting late is felt. A third difficulty is about measurement rather than reach: judgement is assessed during the week, but a national exercise happens once every few years and never twice the same way, so interpretation cannot easily be compared or tracked. Immersive training suits all three, since it is the only setting in which the noise around a cue, the number of readings it allows, its spread across roles and the cost of both kinds of error can be set deliberately, varied, and run again.

The five principles offered here are proposals about design rather than results, and the case rests on public documents and on the account of one of the exercise's organisers. Three lines of work follow. The first is an experiment: compare how quickly and how accurately teams notice and escalate when the level of noise is varied on purpose. The second is to add a weak-signal module to the preparation phase of a national exercise cycle and evaluate it against the indicators in Table 4. The third costs almost nothing, because the assessment machinery already exists: capture, in a standard form, the reasoning that assessors already collect, so that interpretation can be compared between organisations and followed from one cycle to the next. The lynx, as the organisers of ILVES point out, is known by its tracks. Reading tracks is a skill, and skills can be taught.

Disclosure of Interests. The author was one of the organisers of Exercise ILVES 2026 and a member of the team that ran it, as described in Sect. 5. The author has no other competing interests to declare that are relevant to the content of this article.

References

1. Gould, S.: Leading Through Futures Thinking: An Introduction to Using Futures Thinking Questioning and Methods. National Disability Services Queensland, Brisbane (2014)
2. Inayatullah, S.: Six pillars: futures thinking for transforming. Foresight 10(1), 4–21 (2008)
3. Zschau, J., Küppers, A.N. (eds.): Early Warning Systems for Natural Disaster Reduction. Springer, Berlin (2003)
4. Woocher, L.: The effects of cognitive biases on early warning and response. Paper presented at the 49th Annual Convention of the International Studies Association, San Francisco (2008)
5. Lesca, H., Lesca, N.: Strategic Decisions and Weak Signals: Anticipation for Decision-Making. ISTE Ltd and John Wiley & Sons, London (2014)
6. Ansoff, H.I.: Managing strategic surprise by response to weak signals. California Management Review 18(2), 21–33 (1975)
7. Ilmola, L., Kuusi, O.: Filters of weak signals hinder foresight: monitoring weak signals efficiently in corporate decision-making. Futures 38(8), 908–924 (2006)
8. Kuosa, T.: Different approaches of pattern management and strategic intelligence. Technological Forecasting and Social Change 78(3), 458–467 (2011)

9. Schoemaker, P.J.H., Day, G.S., Snyder, S.A.: Integrating organizational networks, weak signals, strategic radars and scenario planning. Technological Forecasting and Social Change 80(4), 815–824 (2013)
10. Hiltunen, E.: The future sign and its three dimensions. Futures 40(3), 247–260 (2008)
11. Lesca, H., Lesca, N.: Weak Signals for Strategic Intelligence: Anticipation Tool for Managers. ISTE Ltd and John Wiley & Sons, London (2011)
12. Bishop, P.: Why is horizon scanning so hard? In: Low, A. (ed.) Decisions in a Complex World – Building Foresight Capabilities, pp. 75–92. National Security Coordination Secretariat, Singapore (2010)
13. Low, A. (ed.): Decisions in a Complex World – Building Foresight Capabilities. National Security Coordination Secretariat, Singapore (2010)
14. Endsley, M.R.: Toward a theory of situation awareness in dynamic systems. Human Factors 37(1), 32–64 (1995)
15. Weick, K.E., Sutcliffe, K.M.: Managing the Unexpected: Sustained Performance in a Complex World. 3rd edn. John Wiley & Sons, Hoboken (2015)
16. Heuer, R.J.: Psychology of Intelligence Analysis. Center for the Study of Intelligence, Central Intelligence Agency, Washington, DC (1999)
17. Kahneman, D., Tversky, A.: Prospect theory: an analysis of decision under risk. Econometrica 47(2), 263–291 (1979)
18. Fischhoff, B.: Hindsight ≠ foresight: the effect of outcome knowledge on judgment under uncertainty. Journal of Experimental Psychology: Human Perception and Performance 1(3), 288–299 (1975)
19. Kolb, D.A.: Experiential Learning: Experience as the Source of Learning and Development. 2nd edn. Pearson FT Press, Upper Saddle River (2014)
20. Makransky, G., Petersen, G.B.: The cognitive affective model of immersive learning (CAMIL): a theoretical research-based model of learning in immersive virtual reality. Educational Psychology Review 33, 937–958 (2021)
21. Tang, X., Chik, N.A., Tang, L.: Optimizing immersive learning in crisis management: cognitive and psychological mechanisms of VR training. Frontiers in Psychology 16, 1695101 (2026)
22. Evain, A., Kropczynski, J., Halse, S., Fertier, A., Panzoli, D., Bénaben, F.: Immersive crisis training: features, observations and the shift from traditional crisis exercises. Journal of Contingencies and Crisis Management 33, e70049 (2025)
23. Harris, D.J., Arthur, T., Kearse, J., Olonilua, M., Hassan, E.K., De Burgh, T.C., et al.: Exploring the role of virtual reality in military decision training. Frontiers in Virtual Reality 4, 1165030 (2023)
24. Juliano, J.M., Schweighofer, N., Liew, S.-L.: Increased cognitive load in immersive virtual reality during visuomotor adaptation is associated with decreased long-term retention and context transfer. Journal of NeuroEngineering and Rehabilitation 19, 106 (2022)
25. Sweller, J.: Cognitive load during problem solving: effects on learning. Cognitive Science 12(2), 257–285 (1988)
26. Government Office of Estonia: Suurõppus ILVES 2026 [ILVES 2026 comprehensive defence exercise]. https://www.riigikantselei.ee/ilves (2026)
27. Government of the Republic of Estonia: ILVES 2026 becomes the largest crisis exercise in Estonia's history. Press release, 12 June 2026. https://valitsus.ee/en
28. Government Office of Estonia: Government Office announces ILVES 2026, one of Estonia's largest crisis exercises. Press release, 8 April 2026. https://www.riigikantselei.ee/en
29. Government Office of Estonia: Kriisiolukorra ja riigikaitse seadus [Crisis Situation and National Defence Act]. https://www.riigikantselei.ee (2026)

30. Kehtna Rural Municipality: Kehtna vald osales riiklikul kriisiõppusel ILVES 2026 [Kehtna municipality took part in the national crisis exercise ILVES 2026] (2026)
31. Luunja Rural Municipality: Õppus ILVES 08–12.06.2026 [Exercise ILVES 8–12 June 2026] (2026)
32. Saku Rural Municipality: ILVES 2026 aitas hinnata meie valla kriisivalmidust [ILVES 2026 helped assess our municipality's crisis preparedness] (2026)
33. Hsieh, H.-F., Shannon, S.E.: Three approaches to qualitative content analysis. Qualitative Health Research 15(9), 1277–1288 (2005)
34. Yin, R.K.: Case Study Research and Applications: Design and Methods. 6th edn. SAGE, Thousand Oaks (2018)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

