Neuroergonomics & Cognitive Engineering, Vol. 207, 2026, 183–190
https://doi.org/10.54941/ahfe1007509

A Cognitive Engineering Approach to Capture the Context of Decision Making for Emergency Call Handlers

Sam Hepenstal[1] and Lewis Lincoln-Gordon[2]

[1]Defence Science and Technology Laboratory (Dstl), Porton Down, Salisbury, UK
[2]Police AI, Milton Keynes, UK

ABSTRACT

Effective decision making is critical for policing, frequently involving high-stakes situations with meaningful consequences for the public, and spanning a wide range of contexts — from officers attending incidents to analysts processing complex data. Whilst it is well recognised that there is a need for sufficient human oversight and accountability when using AI technologies in policing, there is a deeper challenge in designing decision support that genuinely enhances rather than undermines expert performance. Policing decisions often involve experts in their domain whose tacit knowledge, contextual reasoning, and accumulated experience are central to the quality of the decisions they make. When technology is designed primarily to reduce friction and offload cognitive effort, it risks eliminating the very thinking processes that produce good outcomes — sometimes manual effort carries significant cognitive value, and in these situations indiscriminate automation may erode expert performance rather than improve it. This paper demonstrates an applied example of a Cognitive Engineering approach for designing decision support that complements human cognition, through three interconnected steps: first, understanding and modelling the nature of user expertise and cognitive work by mapping workflows; second, discovering the relationships between cognitive activity and overall performance, including where cognitive load represents meaningful value versus unnecessary burden; and third, translating these insights into potential interjection points for AI technology. Taken together, these steps suggest a different way to run the 'discovery' period for AI-enabled decision support: rather than starting from candidate technologies or automation opportunities, discovery begins with characterising expertise in context, then testing which cognitive demands are performance-critical versus merely effortful, and only then forming hypotheses about where AI should interject to amplify expert judgement without eroding it. This methodology was applied to emergency call handling within a policing context, and the findings help challenge assumptions about the potential impact of technological intervention, illustrate the importance of understanding expert cognition before deploying technology, and offer a replicable and applied approach for ensuring AI amplifies rather than replaces human decision-making capability.

Keywords: Cognitive engineering, Decision support, Policing, Artificial intelligence, Call handling, Cognitive task analysis, Human factors

Received 5/8/2026; Revised 6/12/2026; Accepted 6/28/2026; Available online *7/20/2026*

INTRODUCTION AND BACKGROUND

Decision Support in Policing

Across all areas of policing there are a diverse range of decisions being made by humans, from officers attending incidents to staff processing complex data. These decisions are potentially high-risk and high-consequence, carrying a meaningful impact on public safety. For example, when deciding whether to arrest an individual or to prioritise a response and direct resources to a particular incident, a wrong decision can have serious repercussions. In police force control rooms, call handlers manage critical incidents in turbulent and high-pressure environments. Decision support is therefore a critical area for technology development, representing a significant opportunity to help police officers and staff make better and faster decisions.

Decision support involves automating aspects of the decision-making process — for example, to process relevant information about a situation and generate recommendations about suitable responses. Decision support systems could substantially reduce the amount of manual effort required to complete a task, saving significant time and reshaping current workflows. Artificial Intelligence (AI) technologies have the potential to automate decision making in tasks and reduce the friction and effort required to deliver outputs. Due to the nature of decision making in policing and the need for accountability, however, there are necessary constraints on technology that automates or reduces human involvement. The AI Covenant (NPCC, 2023) sets out key guiding principles for the use of AI in policing, underpinned by the notion of human oversight to ensure responsible and accountable use. In the context of decision support, this is typically described as maintaining a 'human-in-the-loop'. However, ensuring a human is nominally in the loop is insufficient if the design of the technology does not genuinely engage and align with the expertise of those humans.

The Context of Decision Making

In policing contexts, decision makers are often experts in their domain — be they analysts, investigators, operational commanders, or call handlers — and this expertise fundamentally informs their thinking, reasoning, and judgements, particularly when dealing with uncertain or ambiguous situations. Many expert-driven activities are time consuming, repetitive, and manual, and it is compelling to introduce systems that automate these processes to reach decisions faster, under the assumption that this leads to better outcomes. However, this assumption warrants scrutiny. If the process of thinking is itself what enables good decisions, then some manual effort may be valuable rather than an unnecessary hindrance. If the aim is to build AI systems that reduce friction, changing current workflows and offloading human cognition, there is a risk of removing the very processes — such as careful, expert-driven reasoning — that enable good outcomes. The challenge, therefore, is not simply to automate, but to understand where automation adds value and where human expertise must remain central. This challenge reflects the nature of complex socio-technical systems, in which the integration of human and machine components must be carefully designed and evaluated. To design technologies that effectively complement how experts think, it is necessary first to understand how people make decisions, the expertise they draw upon, and how that expertise is applied. These aspects can be collectively described as the context of decision making.

HOW CAN WE UNDERSTAND THE CONTEXT OF DECISION MAKING?

Cognitive Engineering: Understanding How People Think

Cognitive Engineering is a user-centred discipline concerned with understanding how people think, reason, and apply expertise during complex problem solving. It provides a basis for designing systems that support rather than supplant human cognition. Throughout complex workflows such as emergency call handling, there are various opportunities where tooling can help users to reduce the cognitive demands of a task — a process known as cognitive offloading (Risko & Gilbert, 2016). While there are tasks that can be offloaded which improve an overall workflow or outputs, research has highlighted meaningful risks to critical thinking, understanding, and memory when cognitive offloading is applied indiscriminately (Sewak, 2025a; Sewak, 2025b; Gerlich, 2025). It is therefore important to consider, prior to seeking to offload a task, whether the process of thinking on behalf of the user is important. If it is, care should be taken to augment and amplify the user's ability to think, rather than to remove that thinking entirely.

Three Steps to Designing Aligned Decision Support

A Cognitive Engineering approach (Hepenstal, 2025) is applied in this paper following three interconnected steps, each building on the last to translate an understanding of expert cognition into grounded system design recommendations.

The first step is to understand the user problem. This involves providing a detailed description of the problem domain, the workflow through which tasks are accomplished, and the cognitive strategies applied by users at different decision points. This step produces a rich picture of what users actually do and how they think.

The second step is to find key relationships between the elements of the task environment — including the information available, the actions taken, the decisions made, and the expertise drawn upon — and the outcomes produced. Critically, this phase distinguishes between cognitive pain points, where mental effort is expended with little benefit to the quality of the output, and cognitive value points, where human thinking meaningfully contributes to performance. This distinction is central to identifying which aspects of a task are appropriate candidates for automation and which require human expertise to be preserved and supported.

The third step is to map these relationships to system representations. The outputs of the preceding steps are used to design user interfaces, information displays, and system behaviours that enable users to maximise the use of their expertise. Recommendations are grounded in the actual cognitive work of practitioners, rather than driven by the availability or novelty of particular technologies.

IMPLEMENTATION: APPLIED TO EMERGENCY CALL HANDLING

Participants and Context

Individual interviews were conducted with four participants from West Midlands Police. Participants held many years of experience across the full range of call handling roles, including receiving 999 and 101 calls, responding to online communications such as email and web chat, and undertaking offline duties such as following up on abandoned calls. The study aimed to produce a detailed account of the call handling workflow and the role of expertise within it.

Methodology

Interviews applied the Critical Decision Method (CDM), a form of Cognitive Task Analysis (Klein et al., 1989), to identify and explore key decision points and the thinking involved when a participant performs a memorable task. A structured script was followed to ensure consistency across interviews. Participants were first asked to describe their roles and the types of tasks they perform day-to-day. They were then asked to describe a particular example call in detail from start to completion. Decision points were identified and discussed, including the thinking and expertise required to deal with uncertainties, any recognised challenges, and the subsequent steps taken by the call handler.

The analytical method broadly followed Emergent Themes Analysis (Wong & Blandford, 2002; Wong, 2004), in which broad themes were identified, indexed, and collated. Interviews were transcribed and utterance data was collated and arranged diagrammatically, clustering utterances that described similar phases of the workflow as broad themes and creating labels to capture the cognition involved — including where call handlers used expertise to manage uncertainty. The diagram was iteratively reviewed and utterances regrouped during analysis, allowing themes to emerge. A final diagram was produced and validated with participants, then used to identify potential technology interjection points.

Initial Analysis: Cognitive Pain and Value Points

A central objective of the analysis was to distinguish between cognitive pain and cognitive value across the call handling workflow. Cognitive pain points are aspects of the task that place a significant burden on the call handler but contribute little to the quality of their decision making — these are appropriate candidates for technological intervention. Cognitive value points are aspects where the call handler's thinking and judgement meaningfully contribute to outcomes — here, it is integral to policing outcomes for the role of the human to be preserved, and technology designed to augment and amplify rather than replace their cognition. This distinction was operationalised through identification of uncertainties and ambiguities, which necessitated call handler expertise. Where the decision-making context involves high uncertainty, expert human reasoning should be augmented; where uncertainty is low, there is greater potential for automation. For example, a call handler must determine the most suitable log type and provide a title to describe the context of the call — there are a range of pre-determined log types to choose from and ambiguity around what to choose to describe the call, with overlaps between categories. The log type determines what form the call handler needs to complete, reflecting the questions they need to ask the caller. Call handlers made various statements which fell within the broad 'log type definition' theme, including:

"We have a lot of log types, and you might just not have used one of them - you don't know it is there…" (Interview 2)

"... one that always catches people out, for example, is when we have traveller encampments, we have a log type for unlawful encampment, but half the people in the room won't remember." (Interview 2)

"If someone has mentioned that they've been beaten up, that's going to be an assault. If someone has mentioned though, it's my brother, my mother and my father, more than likely as long as above 16, it's going to be a domestic." (Interview 4)

"You've got robbery and theft other for your trackable assets, misper so missing persons that's going to create a missing person form for you, DA-domestic abuse… there's quite a few (forms) to be honest… If you click on the crime log type, so theft criminal damage, you'll get a crime desk form." (Interview 4)

These statements indicate cognitive load involved with call handlers having to remember and recall log type options, which can be considered a pain point — for example, that '*half the people in the room won't remember*' a certain log type, highlighting the need for call handlers to recall the various options and the potential for error. There are various log types to consider which may not be familiar to a call handler, despite being pre-determined and thus certain. There is scope for technological intervention to standardise log types and to support surfacing log type options which would reduce cognitive load, whilst improving the quality of the decision. The example statements above also capture valuable cognition where expertise enables a call handler to handle uncertainty around log type definition — the log type can be ambiguous or subjective, whilst being important to inform subsequent questions and inquiries. A call handler utilises their expertise to think about the right definition with appreciation of context.

Results and Implications

All four interviewees described the same general workflow for call handling. A call handler is assigned a duty type by the Service Delivery Officer, including receiving 999 or 101 calls, responding to online communications, or undertaking offline duties. Upon receiving a call, the handler begins a conversation, attempts to understand the situation, and assigns a log type and title. A series of checks are performed to build a picture of the caller and any related history. The call handler applies the THRIVE risk assessment model (Threat, Harm, Risk, Investigation, Vulnerability, Engagement), to grade the call from five options ranging from a 'Priority 1 Emergency' to 'Contact Resolution'. The incident is transferred to dispatch, and the call handler moves on.

The analysis challenged some expectations about where technological intervention would be most impactful. Two contrasting examples illustrate this clearly. First, call handler 'recognition of emergencies', which may have presented a clear use case for technological intervention, was found to be typically obvious to experienced call handlers, with diagnoses within the first minute based on minimal information and tacit expertise.

"...there are some exceptions, but I'd say within the first minute, you know with experience what that grade's going to be (in emergencies) ... you tend to have a feeling for it." (Interview 1)

"Most emergencies… they are quite obvious." (Interview 3)

"(I knew to dispatch as 'Priority 1 Emergency' immediately because) the shouting… someone's obviously gone on their phone, called 999, and then it's all kicked off. (Interview 2)

Automation of emergency diagnosis would therefore yield limited benefit in most cases, due to the cognitive value experts can bring to bear. A call handler also needs to audit and justify their decision making, so transparency and explainability around emergency diagnosis could be an issue. Less obvious emergencies, where subtle cues must be detected, represent a more nuanced opportunity — but technological support here must be guided closely by expert knowledge of what cues are meaningful in context. Second, corroboration of caller location was identified as a high-impact area for intervention. Accurately locating the caller is a priority for all calls, and where location data has low accuracy, call handlers manually search multiple systems. This process is time-consuming and distracts from supporting the caller emotionally. There is low uncertainty in the methods available for corroborating location — the relevant data sources are well-defined — meaning the cognitive demand is largely effort rather than expertise. Automating location corroboration, whilst ensuring the call handler can review the results, represents a high-value intervention.

Based on the analysis of cognitive pain and value points across the workflow, five priority areas for technology intervention were identified:

1. Support for log type classification and title: Technology that helps call handlers define the right log type for a call, provides a clearly formatted title and brief details aligned with best practice.
2. Call pre-processing and real-time triage support: Technology that assists in diagnosing less obvious emergencies and supports automatic population of log information by extracting key entity details from caller audio.

3. Automated corroboration of location information: Technology that automatically fuses address data linked to the caller across available data sources.
4. Tradecraft capture and dynamic information fusion: Technology that learns patterns in the searches performed by call handlers across different situations, automating the fusion of relevant information.
5. Support for form completion, including THRIVE assessments: Technology that helps ensure consistency of inputs and enables review against best practice, with greater automation for lower-risk incidents.

GENERAL IMPLICATIONS AND NEXT STEPS

This paper has demonstrated the application of a structured Cognitive Engineering methodology to the design of decision support for emergency call handlers. The study has some important limitations — the sample comprises only four participants from a single force area, which limits the generalisability of the findings. It is recommended that Cognitive Task Analysis interviews are extended to call handlers from a range of force areas to identify common pain points suited to national solutions. The interviews were intentionally held with experienced call handlers, in order to understand the role of their expertise, however it is recognised that staff retention in control rooms is a challenge and thus technological interventions could be particularly beneficial for less experienced staff.

The study demonstrates a practical method for how we can better understand user thinking and decision making to align AI interventions. The paper underlines the critical importance of understanding the workflow of users, the types of decisions and the thinking and expertise involved in decision making — which enables good decisions — before deploying technology. Expert cognitive work cannot be automated or made frictionless without first developing a clear understanding of the cognitive pain and value embedded within existing workflows. Technology that ignores this risks degrading the very capabilities it was intended to support. The case of emergency recognition in call handling is instructive: a technically ambitious AI system for automatic emergency diagnosis would provide limited benefit because experienced call handlers already excel at this task.

The three-step methodology presented here — understand the user problem, find the key relationships, map relationships to system representations — constitutes a replicable and practical blueprint for understanding the context of decision making in any domain. Whilst this paper has demonstrated its application to emergency call handling, the approach is domain-agnostic and can be applied wherever expert decision making is the subject of technological intervention. The findings also provide important recognition of the limits of cognitive offloading and frictionless design as universal goals. Friction is not always an obstacle; in expert-driven decision making, some forms of effortful engagement are cognitively valuable. Design principles that seek only to minimise effort risk producing systems that are locally easier to use but globally detrimental to performance. Understanding where effort carries value — and where it does not — is therefore a foundational design requirement.

COPYRIGHT INFORMATION

REFERENCES

Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. Societies, 15(1), 6.

Hepenstal, S. (2025). Towards Socio-Technical Evaluation for Artificial Intelligence in Policing. In: Degen, H., Ntoa, S. (eds) Artificial Intelligence in HCI. HCII 2025. Lecture Notes in Computer Science(), vol 15819. Springer, Cham. https://doi.org/10.1007/978-3-031-93412-4_20.

Klein, G. A., Calderwood, R., & MacGregor, D. (1989). Critical decision method for eliciting knowledge. IEEE Transactions on Systems, Man, and Cybernetics, 19(3), 462–472. https://doi.org/10.1109/21.31053

NPCC. (2023). Covenant for using Artificial Intelligence (AI) in Policing. Link: https://science.police.uk/site/assets/files/4682/ai_principles_1_1_1.pdf

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. Trends in Cognitive Sciences, 20(9), 676–688.

Sewak, M. (2025a). The Research Imperative: From Cognitive Offloading to Augmentation: A review of the evidence on Generative AI's impact on our minds. Link: https://pub.towardsai.net/the-research-imperative-from-cognitive-offloading-to-augmentation-7d23efd1da67

Sewak, M. (2025b). A 47% Collapse in Brain Activity: Synthesizing the Data on How LLMs Affect Neural Engagement and Memory. Link: https:/ medium.com/data-science-collective/a-47-collapse-in-brain-activity-9cdf352f349e

Wong, B. L. W., & Blandford, A. (2002). Analysing ambulance dispatcher decision making: Trialing Emergent Themes Analysis. In F. Vetere, L. Johnston, & R. Kushinsky (Eds.), Human Factors 2002 HF2002, the 1st Joint Conference of the Ergonomics Society of Australia (ESA) and the Australian Computer Human Interaction Special Interest Group (OzCHI) (pp. CD–ROM publication).

Wong, B. L. W. (2004). Critical Decision Method Data Analysis. In D. Diaper & N. Stanton (Eds.), Task Analysis for Human-Computer Interaction (pp. 327–346). Lawrence Erlbaum Associates.