

Selecting for Safety: The Academic Evidence Base for a Competency-Based Behavioural Assessment Framework for Control Room Operator Recruitment

Chris Dreyfus-Gibson, Managing Director

Abstract

Control room operators in security, safety, and emergency services environments are required to perform sustained, high-stakes cognitive and communicative work under conditions that most traditional personnel selection methods are poorly equipped to assess. Decades of meta-analytic evidence in personnel psychology consistently demonstrate that unstructured interviews, curriculum vitae reviews, and employment history record checks - the selection methods most commonly relied upon by organisations - possess low criterion-related validity and fail to predict operational performance. This paper presents the theoretical and empirical foundations for a competency-based behavioural assessment framework designed specifically for control room operator selection. Drawing on over sixty peer-reviewed sources across personnel psychology, human factors, safety science, and cognitive ergonomics, it establishes the evidence base for five core competency themes: Foundational Traits, Cognitive and Decision-Making Capabilities, Communication and Coordination, Learning and Adaptability, and Values and Role Fit. For each theme, and for the eighteen individual competencies within them, the paper identifies the primary academic constructs and meta-analytic evidence

supporting their predictive validity in safety-critical and surveillance environments. The paper additionally sets out the methodological rationale for behaviourally anchored, pattern-based assessment as a superior alternative to traditional selection approaches. Taken together, this evidence provides a robust scientific basis for structured competency assessment as the appropriate standard for control room operator recruitment.

Keywords: competency-based assessment, personnel selection, control room operations, behavioural assessment, safety-critical environments, human factors, vigilance, situational awareness, operator performance

1. Introduction

Control rooms present a particular and demanding human performance challenge. Whether in security operations, emergency dispatch, critical national infrastructure protection, or command and control environments, operators are required to sustain attention across extended monitoring periods, detect low-frequency but high-consequence events, manage competing information streams under time pressure, communicate accurately with multiple parties, and make consequential decisions, often with incomplete information and without direct supervision. The operational environment shares critical characteristics with other safety-critical domains - aviation, nuclear power, offshore operations, and process industries - where human error is not merely a performance issue but a risk management concern (Flin, O'Connor, & Crichton, 2008; Reason, 1997).

Despite this, the personnel selection practices used to recruit control room operators have lagged significantly behind the scientific evidence on what actually predicts success in such roles. The majority of organisations continue to rely on methods that the personnel psychology

literature has consistently shown to be weak predictors of operational performance: unstructured interviews, work history reviews, and the assessment of prior experience or qualifications (Schmidt & Hunter, 1998; Sackett, Zhang, Berry, & Lievens, 2022). The consequences of this misalignment are measurable. High training attrition rates, inconsistent operator performance, difficulty sustaining vigilance during low-activity periods, and costly rehiring cycles are well-documented operational problems in control room environments that structured selection could substantially reduce.

This paper addresses the question: what does the scientific evidence tell us about how to select control room operators effectively? In answering this question, it establishes the empirical basis for a framework of five competency themes and eighteen individual competencies, each grounded in peer-reviewed research. It then addresses the methodological question of how those competencies should be assessed, setting out the psychometric rationale for structured, behaviourally anchored assessment as the appropriate approach for high-consequence selection contexts.

The framework described here was developed through integration of three bodies of literature: personnel psychology (particularly meta-analyses of selection method validity); human factors and cognitive ergonomics (particularly research on operator performance in monitoring, vigilance, and decision-making tasks); and safety science (particularly research on non-technical skills and safety culture in high-hazard industries). The convergence across these distinct disciplines on a common set of operationally relevant constructs is itself a strong argument for the framework's scientific credibility.

2. The Failure of Traditional Selection Methods in Safety-Critical Contexts

The most consequential finding in over a century of personnel selection research is both straightforward and persistently ignored in practice: the methods that most organisations use to hire people are among the weakest predictors of how those people will actually perform. Schmidt and Hunter (1998), in the most cited paper in the personnel selection literature, meta-analysed 19 distinct selection procedures across 85 years of accumulated data. They found that unstructured employment interviews achieved a criterion-related validity of $r = .38$, while education and years of job experience - the core content of most curriculum vitae assessments - returned validities of $r = .10$ and $r = .18$ respectively. Classified within a broader composite approach, structured work sample tests ($r = .54$) and structured interviews combined with cognitive ability testing substantially outperformed all traditional methods.

More recently, and more significantly, Sackett, Zhang, Berry, and Lievens (2022) revised the established meta-analytic estimates downward following the identification of a systematic overcorrection for range restriction in prior meta-analyses. Under their corrected analysis, the criterion-related validity of the unstructured interview falls to $r = .19$ - barely above chance as a standalone predictor. This is the most methodologically robust estimate currently available, and it places the unstructured interview on a level comparable to graphology and reference checks, neither of which serious practitioners regard as defensible tools for safety-critical selection.

The problem is compounded by the well-documented tendency of interviewers to overestimate their ability to make accurate judgements in the absence of structure. Levashina, Hartwell, Morgeson, and Campion (2014) found that fewer than a third of practitioners had received any formal interview training, yet confidence in making accurate people assessments remained uniformly high. Campion, Palmer, and Campion (1997) demonstrated through detailed review

that validity increases monotonically with interview structure, identifying fifteen specific structural components that each independently improve predictive accuracy, from the use of job-analysis-based questions to the implementation of anchored rating scales. Where none of these components are present, as is typical in informal screening interviews, the assessment is effectively measuring articulacy and impression management rather than the operational traits that control room roles require.

In control room contexts, the cost of selection error extends beyond direct rehiring expense. Research in related safety-critical domains is instructive. Beus, Dhanani, and McCord (2015) demonstrated a significant relationship between personality traits measurable at selection and rates of unsafe workplace behaviour. The Travelers (2024) Injury Impact Report found that employees are four to six times more likely to be involved in a serious incident during their first month in post, with 35% of all injuries occurring within the first year of employment - a period during which the behavioural fit of the individual to the role is most consequential. In high-governance control room environments, where the consequences of an operator's failure to detect an alarm, miscommunicate an instruction, or make a procedurally incorrect decision may cascade into major operational or safety incidents, the argument for structured, evidence-based selection is not merely commercial. It is a duty of care.

3. The Case for Competency-Based Behavioural Assessment

The scientific consensus in personnel psychology points clearly towards structured assessment of specific, job-relevant competencies as the most valid and equitable approach to personnel selection. Arthur, Day, McNelly, and Edens (2003), in their meta-analysis of assessment centre dimensions across 168 studies, found criterion-related validities of between .25 and .39 for specific behavioural dimensions - notably exceeding holistic overall assessment ratings - with

problem-solving ability returning the highest observed validity ($\rho = .39$). Their regression modelling demonstrated that a composite of four behavioural dimensions explained approximately 20% of performance variance, providing substantially greater predictive value than subjective global impressions.

The superiority of structured behavioural assessment derives from several converging mechanisms. First, by anchoring evaluation to specific, operationally defined behaviours rather than impressionistic overall ratings, it reduces the construct-irrelevant variance that inflates and distorts unstructured assessments. Smith and Kendall (1963), who introduced Behaviourally Anchored Rating Scales (BARS) in the personnel psychology literature, found inter-rater reliabilities exceeding .97 when specific behavioural examples were used to define points on evaluative continua. Second, multi-method assessment allows for the observation of behaviours across diverse conditions that more closely mirror the demands of the operational role - addressing the fidelity limitation of single-method approaches (Lievens & Patterson, 2011). Third, pattern-based assessment, which looks for consistent behavioural themes across multiple exercises rather than accepting or rejecting candidates on the basis of single-task performance, better captures the stable, dispositional characteristics that predict long-term operational reliability.

Situational judgement tests (SJTs), which present candidates with realistic scenarios and assess the quality and appropriateness of their decision-making responses, have demonstrated generalizable criterion-related validity of $\rho = .34$ (McDaniel, Morgeson, Finnegan, Campion, & Braverman, 2001), with incremental validity beyond cognitive ability and personality measures. Christian, Edwards, and Bradley (2010) further demonstrated that SJTs assessing teamwork and interpersonal skills predicted contextual performance outcomes, and that video-

based SJTs outperformed text-based versions, supporting the use of behaviourally realistic scenarios.

In safety-critical industries specifically, there is strong institutional endorsement for competency-based selection. The International Atomic Energy Agency (IAEA, 2020) recommends structured behavioural competency assessment for nuclear operations staff throughout the employment lifecycle, grounded in a risk-informed analysis of role demands. The UK Health and Safety Executive (HSE, 2003) established 15 key principles for competence management in high-hazard industries and noted that inadequate competence assurance had been a contributing factor in several major industrial disasters. In control room environments, where the consequences of poor selection decisions are expressed not in reduced productivity but in risk to people, assets, and critical systems, this authoritative endorsement for structured competency assessment carries direct operational weight.

4. Foundational Traits: The Dispositional Substrate of Reliable Performance

The first competency theme addresses traits that are difficult to train but highly predictive of reliability and trustworthiness in operational roles. These are constructs that exist at the level of personality, character, and individual difference - they are relatively stable across time and context, and they determine whether an operator can be trusted to behave consistently, accurately, and honestly when unsupervised.

Sustained Attention and Vigilance

The vigilance decrement - the documented deterioration in signal detection performance over time in continuous monitoring tasks - is one of the most replicated phenomena in applied

cognitive psychology. See, Howe, Warm, and Dember (1995), in their meta-analysis of vigilance research, demonstrated that the decrement is far more prevalent and more severe than previously acknowledged, particularly as task difficulty increases and events are non-successive. Warm, Parasuraman, and Matthews (2008) challenged the prevailing assumption that vigilance is passive and low-demand, providing convergent behavioural, neuroimaging, and subjective workload evidence that sustained monitoring tasks impose high cognitive workload and deplete attentional resources, with performance deterioration beginning within as few as fifteen to thirty minutes.

Research on CCTV operators specifically confirms that these effects apply directly to security surveillance contexts. Donald, Donald, and Thatcher (2015) found that control room specialists detected only 50% of target behaviours across a monitoring session, with vigilance decrements differentiating novice from experienced personnel within the first hour. Donald (2019) proposed a model of CCTV operator performance in which expectancy and situational awareness mediate attentional orientation, and recommended selection approaches that identify individuals with greater intrinsic capacity for sustained attention. Individual differences in sustained attention - measurable at assessment - are therefore a meaningful and consequential selection criterion.

Conscientiousness and Reliability

Barrick and Mount (1991) meta-analysed 117 studies (N = 23,994) examining the relationship between Big Five personality dimensions and job performance. Conscientiousness was the only dimension to show consistent, generalisable validity across all performance criteria and all occupational groups. In safety-critical contexts specifically, Christian, Bradley, Wallace, and Burke (2009) used meta-analytic path modelling to confirm that conscientiousness and emotional stability predict safety performance through their effects on safety knowledge and

safety motivation. Beus et al. (2015), in a meta-analysis of 69 studies, further confirmed that conscientiousness is negatively associated with unsafe behaviours, independent of situational factors. Control rooms depend on dependable, repeatable behaviour rather than improvisation, and conscientiousness is precisely the dispositional construct that predicts this.

Emotional Regulation

Gross (1998) introduced the process model of emotion regulation, distinguishing between antecedent-focused strategies (particularly cognitive reappraisal) and response-focused strategies (particularly expressive suppression). Reappraisal reduces both subjective emotional experience and expression without increasing physiological arousal, while suppression reduces expression but simultaneously increases cardiovascular activation - with direct consequences for the quality of sustained cognitive performance. LeBlanc (2009) demonstrated that acute stress impairs divided attention, working memory, information retrieval, and decision-making even in experienced professionals, following a pattern consistent with the Yerkes-Dodson law. Operators who cannot regulate emotional responses to stressful events risk compounding operational pressure through degraded communication, impaired judgement, and inadvertent escalation. The capacity to remain calm and purposeful under pressure is not a personality luxury; it is an operational necessity.

Integrity, Trustworthiness, and Accountability

Ones, Viswesvaran, and Schmidt (1993) conducted the most comprehensive meta-analysis of integrity test validity, aggregating 665 coefficients across 576,460 participants. Mean operational predictive validity for overall job performance was $\rho = .41$, among the strongest effects in the personnel selection literature. Integrity tests additionally predicted counterproductive work behaviours including theft, disciplinary problems, and absenteeism. Sackett and Schmitt (2012) confirmed that under predictive study designs, with applicant

samples and non-self-report criteria - the conditions most directly relevant to actual selection decisions - integrity measures remain robust predictors. In control room environments, where operators are trusted with sensitive systems, confidential information, and consequential decisions that may not be directly observable, the predictive value of pre-selection integrity assessment is self-evidently significant.

Accountability - the willingness to own decisions, acknowledge errors transparently, and engage in corrective learning - is supported by a separate body of evidence relating to safety culture. Edmondson (1999) established in 51 work teams that psychological safety, which is partly determined by individuals' willingness to take accountability, was significantly associated with team learning behaviour and, through it, team performance. Reason (1997) and Dekker (2017) both identified accountability culture as a foundational component of high-reliability safety systems: organisations that cannot generate honest incident reports because operators fear blame cannot learn from their errors and are systematically blind to accumulating risk.

5. Cognitive and Decision-Making Competencies

The second competency theme addresses how operators think under uncertainty - the cognitive processes through which they interpret information, form situational assessments, and commit to action. These competencies are particularly relevant in control room environments because, as Flin, O'Connor, and Crichton (2008) note, up to 80% of accident causes are attributable to human factors, and cognitive errors in interpretation, prioritisation, and decision-making are among the most common contributing mechanisms.

Situational Judgement

Rasmussen (1983) provided the foundational cognitive model for understanding operator behaviour in complex systems through the Skills, Rules, and Knowledge (SRK) framework. He identified three levels of cognitive control: skill-based behaviour (automated, requiring minimal attentional resource), rule-based behaviour (IF-THEN procedural, moderate resource demand), and knowledge-based behaviour (deliberate analytical reasoning, high resource demand and high error probability). Chen, Li, and Zio (2023) quantified human error probabilities across these levels using three error databases, confirming that knowledge-based processing carries the highest error likelihood. Critically, time pressure causes regression from rule-based towards knowledge-based processing, increasing error rates, while training and practice enable the reverse transition.

Klein (1993, 1998) extended this understanding through naturalistic decision-making research, finding that experienced decision-makers in high-stakes environments typically recognise situational patterns and mentally simulate a single course of action, selecting a workable option as their first consideration in approximately 80% of decisions. This recognition-primed decision (RPD) model explains why pattern-familiarisation and situational context are so important to effective operator performance, and why assessment of contextual judgement - rather than only procedural knowledge - provides more valid predictions of real-world performance.

Pattern Recognition

Green and Swets (1966) established Signal Detection Theory (SDT) as the framework for understanding how observers detect signals against a background of noise, separating genuine sensitivity (d') from response bias. In control room contexts, operators must detect anomalous events from among a continuous stream of routine data; SDT provides the

principled framework for evaluating this capability and for understanding the distinction between operators who genuinely detect more signals and those who respond more liberally at the cost of higher false alarm rates. Chase and Simon (1973) demonstrated that expert pattern recognition is domain-specific and is built from approximately 50,000 familiar configurations or ‘chunks’ of domain knowledge in chess masters - a finding that has been extensively replicated in other expert domains and that implies the value of assessing recognition-based cognitive processing capacity at selection, before domain-specific chunks are acquired through training.

Prioritisation, Task Switching, and Risk Awareness

Wickens (2002, 2008) developed the Multiple Resources Theory of attention, demonstrating that interference between concurrent tasks depends on whether they draw on shared or separate processing stages, sensory modalities, processing codes, and visual channels. Control room operators routinely manage multiple simultaneous demands - alarm monitoring, radio communication, camera review, incident logging, team coordination - and the capacity to allocate and reallocate attentional resources efficiently is therefore a fundamental performance requirement. Monsell (2003) established that task-switching costs - the performance decrements that occur when cognitive sets must be shifted between tasks - are never fully eliminated by preparation time, implying that some attentional capacity overhead is inherent in multi-task environments regardless of operator competence. Selection must therefore identify individuals whose cognitive resources are sufficient to absorb this inherent cost without consequential performance degradation.

Decision confidence and risk calibration are addressed by Tversky and Kahneman (1974) and the body of heuristics and biases research: operators who rely on availability, representativeness, or anchoring heuristics in risk assessment consistently produce

systematically biased probability estimates. Both overconfidence and under-confidence are identified as failure modes in time-sensitive environments - the former leading to premature commitment and the latter to costly hesitation. Slovic's (1987) psychometric work on risk perception further demonstrated that laypeople's risk judgements are driven by the perceived characteristics of hazards rather than probabilistic reasoning - a finding that underlines the importance of assessing calibrated, structured risk awareness as distinct from general intelligence.

6. Communication and Coordination

The third competency theme addresses how operators interact within the operational system. Communication failures are not merely a contributing factor in operational errors - they are frequently the primary mechanism. Tyler and colleagues (2025), in a systematic review of 46 studies covering 67,826 patients, found that poor communication was the sole cause of safety incidents in over 10% of cases and contributed to approximately 40% of adverse events. In control room and command and control environments, where operators function as distributed nodes within an integrated information system rather than as autonomous individual performers, the quality of communication is inseparable from the quality of operational outcomes.

Helmreich, Merritt, and Wilhelm (1999) traced the development of Crew Resource Management (CRM) training in commercial aviation to the 1979 NASA workshop at which human error - specifically failures in interpersonal communication, leadership, and decision-making - was identified as the primary cause of the majority of serious aviation accidents. The subsequent widespread adoption of CRM principles across aviation, nuclear operations, surgical teams, and emergency services reflects the breadth of evidence that structured

communication practices reduce incident rates. The competencies that CRM targets - clear structured communication, closed-loop confirmation, assertiveness in raising concerns, and shared situational updating - are precisely those that the Communication and Coordination theme of an operator assessment framework must evaluate.

Endsley's (1995) three-level theory of Situation Awareness - perception of environmental elements, comprehension of their meaning, and projection of future states - has become the dominant framework for understanding shared operational understanding in team environments. DeChurch and Mesmer-Magnus (2010) confirmed in meta-analytic review that teams with greater convergence in shared mental models show significantly better coordination and higher performance, particularly in dynamic, complex tasks. Salas, Sims, and Burke (2005) identified closed-loop communication as one of three coordinating mechanisms underpinning effective teamwork, alongside shared mental models and mutual trust. Kluger and colleagues (in press), in a meta-analysis of 664 effect sizes covering 400,020 observations, established a corrected correlation of $\rho = .44$ between perceived listening quality and work outcomes - confirming that the ability to accurately receive, interpret, and retain information is a measurable and consequential individual competency.

7. Learning and Adaptability

The fourth competency theme addresses how quickly and safely operators develop. Most candidates entering control room roles will be new to the specific operational environment, the systems in use, and the procedural frameworks governing their conduct. Selection therefore cannot rely solely on existing capability - it must identify individuals who will acquire capability effectively and safely during training and the early operational period.

De Meuse (2019), in a meta-analysis of 20 field studies ($N = 4,613$), found a corrected true population correlation of $\rho = .74$ between learning agility and leader performance, and $\rho = .75$ for leader potential - placing learning agility among the most powerful predictors in organisational psychology. De Meuse, Dai, and Hallenbeck (2010) demonstrated that high learning-agile individuals were significantly more likely to be promoted and receive higher performance ratings, and that learning agility predicted future performance beyond current performance levels - precisely the predictive relationship that pre-training selection requires.

The theoretical grounding for procedural thinking as a safety-critical competency rests on Rasmussen's (1983) SRK framework and Reason's (1990, 1997) error models. Reason established that violations - deliberate deviations from specified procedures - represent a categorically different failure mode from errors, and that organisations whose cultures permit or normalise rule-bending progressively erode the procedural defences that prevent simultaneous failure conditions from aligning into major incidents. Control room operators who understand procedures not as bureaucratic constraints but as engineered safeguards, and who can recognise the limits of their authority and the conditions under which escalation is required, are more reliably safe and more easily supervised.

Self-awareness and reflectiveness are supported by Yeung and Summerfield's (2012) metacognition research, which confirmed that individuals vary meaningfully in their ability to monitor their own cognitive processes, detect errors prior to external feedback, and calibrate confidence to objective performance. Schon's (1983) influential work on reflective practice established that practitioners who engage in active reflection during performance - not only after it - are better equipped to detect when routine approaches are insufficient and to adapt their thinking for novel situations. In control room environments, where no two incidents are identical and where operators may need to make critical judgements about when to follow

procedure and when to escalate, metacognitive self-awareness is directly operationally relevant.

8. Values and Role Fit

The fifth competency theme addresses why the person should hold this responsibility. Technical and cognitive capability are necessary but not sufficient conditions for effective control room performance. An operator who possesses appropriate cognitive capacity and communication skill but who is motivated by excitement rather than duty of care, or who resists procedural governance in favour of personal judgement, is a qualitatively different risk from one whose capability is still developing. Values and role fit determine whether an individual's motivation will sustain their commitment to safe, governed behaviour across the full range of conditions they will encounter in post.

Kristof-Brown, Zimmerman, and Johnson (2005), in a meta-analysis of 172 studies covering 836 effect sizes, found that person-organisation fit showed strong relationships with job satisfaction ($\rho = .44$), organisational commitment ($\rho = .51$), and intent to leave ($\rho = -.35$). Cable and Judge (1996, 1997) confirmed that P-O fit is driven primarily by values congruence rather than demographic similarity, and that interviewers can assess applicant-organisation values alignment with meaningful accuracy. In safety-critical contexts, where an operator's sustained commitment to procedural governance and duty of care must remain stable across long shifts, low-stimulation periods, and stressful incidents, values alignment is an operational prerequisite.

Griffin and Neal (2000) established the foundational distinction between safety compliance - adherence to mandated safety procedures - and safety participation - proactive engagement with safety beyond minimum requirements. Both are meaningful performance outcomes, but

safety participation requires genuine internalisation of protective values rather than mere compliance with external monitoring. Neal and Griffin (2006) confirmed longitudinally that safety motivation predicts safety behaviour, which predicts reduced accident rates over time. Zohar (2000) demonstrated that group-level safety climate - substantially influenced by the values and practices of individual team members - predicts microaccident rates independent of other factors. Selecting operators whose protective mindset is intrinsic rather than performative therefore has direct implications for the safety climate of the team they join.

The research on authority gradient and governance compliance is also instructive. Sydor and colleagues (2013) demonstrated in a randomised controlled trial that hierarchy gradients powerfully inhibit speaking up even when patient safety is directly at stake, and that structural changes alone are insufficient to overcome this. Sekar and colleagues (2022), reviewing authority gradient research across healthcare, aviation, rail, and maritime, found that up to a third of personnel have concerns about challenging authority in safety-critical moments. Selecting individuals who value structured escalation pathways, understand hierarchical governance as a functional safeguard rather than a constraint, and are comfortable operating within defined authority limits is therefore a meaningful risk reduction intervention.

9. Assessment Methodology: Pattern-Based, Behaviourally Anchored

The scientific case for competency-based selection is not merely about what to assess, but how to assess it. Three methodological principles, each supported by independent empirical evidence, define the appropriate approach.

First, assessment should be behaviourally grounded rather than impression-based. Smith and Kendall (1963) demonstrated that anchoring evaluative continua with concrete behavioural examples - specific descriptions of what low, moderate, and high performance actually looks

like in practice - dramatically improves inter-rater reliability and reduces the influence of irrelevant impression factors. Campion et al. (1997) identified behaviourally anchored evaluation as one of the most impactful structural components for improving interview validity. The principle extends beyond interviews to all forms of behavioural observation: assessors who are required to record specific behavioural evidence before assigning ratings make substantially more accurate and defensible judgements than those who assign ratings directly.

Second, assessment should be pattern-based rather than pass/fail. No single observation, however well-designed, can reliably represent the full range of behaviours relevant to operational performance. Individuals may underperform on a single task due to momentary inattention, mild anxiety about assessment, or misunderstanding of instructions. Equally, candidates who perform well on a single high-salience task may not sustain that performance across routine conditions. Assessment across multiple scenarios, exercises, and modes - individually completed cognitive tasks, scenario-based discussions, and collaborative group exercises - allows the assessor to identify consistent behavioural themes rather than reacting to isolated peaks or troughs. This is precisely the approach endorsed by Arthur et al. (2003), whose meta-analysis demonstrated that the predictive validity of dimension-level assessment improves when multiple exercises contribute to each dimension rating.

Third, the assessment should be predictive rather than retrospective: it should be designed to identify traits and capabilities that will predict performance in the role, not merely suitability for the assessment itself. This distinction matters because candidates can be coached to perform well on a specific assessment format without the underlying competencies developing commensurately. The most robust defences against this are multi-method assessment (reducing the value of coaching for any single method), competency anchors derived directly from job analysis (ensuring the constructs being measured are operationally relevant), and the use of

timing, accuracy, and consistency data from objective performance tasks alongside structured behavioural observation. Together, these methods produce evidence that is comparable across candidates, defensible under audit, and predictive of the operational behaviours that determine safe and effective control room performance.

10. Conclusion

The scientific case for competency-based behavioural assessment of control room operators is not speculative. It rests on a robust, cross-disciplinary evidence base that has been accumulating for over a century in personnel psychology and several decades in human factors, safety science, and cognitive ergonomics. The evidence converges on a set of findings that are individually well-supported and collectively compelling.

Traditional selection methods, and unstructured interviews in particular, possess criterion-related validity that Sackett and colleagues (2022) have now estimated at $r = .19$ after correcting for longstanding statistical artefacts. They measure articulacy, social confidence, and impression management rather than the operational traits that control room roles demand. Structured, behaviourally anchored competency assessment, by contrast, produces validities of .35 to .45 for individual dimensions, with composite assessments explaining 20% or more of performance variance - a scientifically meaningful and practically significant improvement.

The five competency themes set out in this paper - Foundational Traits, Cognitive and Decision-Making Capabilities, Communication and Coordination, Learning and Adaptability, and Values and Role Fit - each rest on independent empirical foundations. Sustained attention and vigilance are measurably variable between individuals and directly predict control room detection performance. Conscientiousness predicts both safety behaviour and operational reliability across all occupational groups. Emotional regulation determines whether cognitive

performance degrades under the pressure conditions that are routine in high-consequence operations. Integrity and accountability are prerequisites for the trust that safety-critical roles require and for the learning cultures that high-reliability organisations depend upon.

Cognitive competencies including situational judgement, pattern recognition, and prioritisation are grounded in well-established theoretical models of operator cognition that have been validated in control room, aviation, nuclear, and emergency service contexts. Communication competencies are supported by some of the most direct evidence of all: poor communication is a measurable, preventable cause of a substantial proportion of safety incidents across high-consequence environments. Learning agility predicts who will succeed in training; procedural thinking and self-awareness predict who will remain safe and continue to develop once operational. Values alignment determines whether an individual's motivation will sustain compliant, conscientious, and safety-oriented behaviour across the full operational lifecycle.

Taken together, these findings provide the scientific foundation for a structured, validated, and defensible approach to control room operator selection. In environments where an individual's capacity to detect a signal, communicate an instruction, or escalate a concern at the right moment may be the difference between a contained incident and a significant operational failure, selecting on the basis of evidence rather than instinct is not merely good practice. It is the responsible standard.

References

- Arthur, W., Jr., Day, E. A., McNelly, T. L., & Edens, P. S. (2003). A meta-analysis of the criterion-related validity of assessment centre dimensions. *Personnel Psychology*, 56(1), 125–153.
- Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44(1), 1–26.
- Beus, J. M., Dhanani, L. Y., & McCord, M. A. (2015). A meta-analysis of personality and workplace safety: Addressing unanswered questions. *Journal of Applied Psychology*, 100(2), 481–498.
- Cable, D. M., & Judge, T. A. (1996). Person–organisation fit, job choice decisions, and organisational entry. *Organizational Behavior and Human Decision Processes*, 67(3), 294–311.
- Cable, D. M., & Judge, T. A. (1997). Interviewers' perceptions of person–organisation fit and organisational selection decisions. *Journal of Applied Psychology*, 82(4), 546–561.
- Campion, M. A., Palmer, D. K., & Campion, J. E. (1997). A review of structure in the selection interview. *Personnel Psychology*, 50(3), 655–702.
- Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, 4, 55–81.
- Chen, J., Li, Z., & Zio, E. (2023). Quantification of human behaviour levels by extending Rasmussen's SRK model and the effects of time pressure and training on levels switching. *Heliyon*, 9(5), e15359.

- Christian, M. S., Bradley, J. C., Wallace, J. C., & Burke, M. J. (2009). Workplace safety: A meta-analysis of the roles of person and situation factors. *Journal of Applied Psychology*, 94(5), 1103–1127.
- Christian, M. S., Edwards, B. D., & Bradley, J. C. (2010). Situational judgement tests: Constructs assessed and a meta-analysis of their criterion-related validities. *Personnel Psychology*, 63(1), 83–117.
- Davies, D. R., & Parasuraman, R. (1982). *The psychology of vigilance*. Academic Press.
- DeChurch, L. A., & Mesmer-Magnus, J. R. (2010). The cognitive underpinnings of effective teamwork: A meta-analysis. *Journal of Applied Psychology*, 95(1), 32–53.
- De Meuse, K. P. (2019). A meta-analysis of the relationship between learning agility and leader success. *Journal of Organizational Psychology*, 19(1), 25–34.
- De Meuse, K. P., Dai, G., & Hallenbeck, G. S. (2010). Learning agility: A construct whose time has come. *Consulting Psychology Journal: Practice and Research*, 62(2), 119–130.
- Dekker, S. (2017). *Just culture: Restoring trust and accountability in your organisation* (3rd ed.). CRC Press.
- DeRue, D. S., Ashford, S. J., & Myers, C. G. (2012). Learning agility: In search of conceptual clarity and theoretical grounding. *Industrial and Organizational Psychology*, 5(3), 258–279.
- Donald, F. M. (2019). A model of CCTV surveillance operator performance. *Ergonomics*, 62(4), 503–516.

- Donald, F. M., Donald, C., & Thatcher, A. (2015). Work exposure and vigilance decrements in closed-circuit television surveillance. *Applied Ergonomics*, 47, 1–8.
- Edmondson, A. C. (1999). Psychological safety and learning behaviour in work teams. *Administrative Science Quarterly*, 44(2), 350–383.
- Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors*, 37(1), 32–64.
- Fischhoff, B., Slovic, P., Lichtenstein, S., Read, S., & Combs, B. (1978). How safe is safe enough? A psychometric study of attitudes toward technological risks and benefits. *Policy Sciences*, 9, 127–152.
- Flin, R., O'Connor, P., & Crichton, M. (2008). *Safety at the sharp end: A guide to non-technical skills*. Ashgate.
- Gaugler, B. B., Rosenthal, D. B., Thornton, G. C., III, & Bentson, C. (1987). Meta-analysis of assessment centre validity. *Journal of Applied Psychology*, 72(3), 493–511.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. Wiley.
- Griffin, M. A., & Neal, A. (2000). Perceptions of safety at work: A framework for linking safety climate to safety performance, knowledge, and motivation. *Journal of Occupational Health Psychology*, 5(3), 347–358.
- Gross, J. J. (1998). The emerging field of emotion regulation: An integrative review. *Review of General Psychology*, 2(3), 271–299.

- Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Advances in psychology* (Vol. 52, pp. 139–183). North-Holland.
- Health and Safety Executive. (2003). *Competence assessment for the hazardous industries* (Research Report 086). HSE Books.
- Helmreich, R. L., Merritt, A. C., & Wilhelm, J. A. (1999). The evolution of crew resource management training in commercial aviation. *International Journal of Aviation Psychology*, 9(1), 19–32.
- International Atomic Energy Agency. (2020). *Assessing behavioural competencies of employees in nuclear facilities* (TECDOC-1917). IAEA.
- Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge University Press.
- Klein, G. A. (1993). A recognition-primed decision (RPD) model of rapid decision making. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsombok (Eds.), *Decision making in action: Models and methods* (pp. 138–147). Ablex.
- Klein, G. A. (1998). *Sources of power: How people make decisions*. MIT Press.
- Kluger, A. N., & Itzhakov, G. (2022). The power of listening at work. *Annual Review of Organizational Psychology and Organizational Behavior*, 9, 121–146.
- Kristof-Brown, A. L., Zimmerman, R. D., & Johnson, E. C. (2005). Consequences of individuals' fit at work: A meta-analysis of person-job, person-organisation, person-group, and person-supervisor fit. *Personnel Psychology*, 58(2), 281–342.

- LeBlanc, V. R. (2009). The effects of acute stress on performance: Implications for health professions education. *Academic Medicine*, 84(Suppl 10), S25–S33.
- Levashina, J., Hartwell, C. J., Morgeson, F. P., & Campion, M. A. (2014). The structured employment interview: Narrative and quantitative review of the research literature. *Personnel Psychology*, 67(1), 241–293.
- Lievens, F., & Patterson, F. (2011). The validity and incremental validity of knowledge tests, low-fidelity simulations, and high-fidelity simulations for predicting job performance in advanced-level high-stakes selection. *Journal of Applied Psychology*, 96(5), 927–940.
- McDaniel, M. A., Morgeson, F. P., Finnegan, E. B., Campion, M. A., & Braverman, E. P. (2001). Use of situational judgement tests to predict job performance: A clarification of the literature. *Journal of Applied Psychology*, 86(4), 730–740.
- McDaniel, M. A., Hartman, N. S., Whetzel, D. L., & Grubb, W. L., III. (2007). Situational judgement tests, response instructions, and validity: A meta-analysis. *Personnel Psychology*, 60(1), 63–91.
- Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, 7(3), 134–140.
- Neal, A., & Griffin, M. A. (2006). A study of the lagged relationships among safety climate, safety motivation, safety behaviour, and accidents. *Journal of Applied Psychology*, 91(4), 946–953.
- Ones, D. S., Viswesvaran, C., & Schmidt, F. L. (1993). Comprehensive meta-analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance. *Journal of Applied Psychology*, 78(4), 679–703.

- Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-13(3), 257–266.
- Reason, J. (1990). *Human error*. Cambridge University Press.
- Reason, J. (1997). *Managing the risks of organisational accidents*. Ashgate.
- Sackett, P. R., & Schmitt, N. (2012). On reconciling conflicting meta-analytic findings regarding integrity test validity. *Journal of Applied Psychology*, 97(3), 550–556.
- Sackett, P. R., Zhang, C., Berry, C. M., & Lievens, F. (2022). Revisiting meta-analytic estimates of validity in personnel selection: Addressing systematic overcorrection for restriction of range. *Journal of Applied Psychology*, 107(11), 2040–2068.
- Salas, E., Sims, D. E., & Burke, C. S. (2005). Is there a ‘Big Five’ in teamwork? *Small Group Research*, 36(5), 555–599.
- Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124(2), 262–274.
- Schon, D. A. (1983). *The reflective practitioner: How professionals think in action*. Basic Books.
- See, J. E., Howe, S. R., Warm, J. S., & Dember, W. N. (1995). Meta-analysis of the sensitivity decrement in vigilance. *Psychological Bulletin*, 117(2), 230–249.
- Sekar, H., Dharmasena, D., et al. (2022). Understanding authority gradient: Tips for speaking up for patient safety. *The Obstetrician & Gynaecologist*, 24(4), 272–280.

Slovic, P. (1987). Perception of risk. *Science*, 236(4799), 280–285.

Smith, P. C., & Kendall, L. M. (1963). Retranslation of expectations: An approach to the construction of unambiguous anchors for rating scales. *Journal of Applied Psychology*, 47(2), 149–155.

Sydor, D. T., Bould, M. D., Naik, V. N., et al. (2013). Challenging authority during a life-threatening crisis: The effect of operating theatre hierarchy. *British Journal of Anaesthesia*, 110(3), 463–471.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124–1131.

Tyler, N., Gorelik, A., Hodkinson, A., et al. (2025). Impacts of communication type and quality on patient safety incidents: A systematic review. *Annals of Internal Medicine*, 178(5).

Warm, J. S., Parasuraman, R., & Matthews, G. (2008). Vigilance requires hard mental work and is stressful. *Human Factors*, 50(3), 433–441.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, 3(2), 159–177.

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors*, 50(3), 449–455.

Yeung, N., & Summerfield, C. (2012). Metacognition in human decision-making: Confidence and error monitoring. *Philosophical Transactions of the Royal Society B*, 367(1594), 1310–1321.

Zohar, D. (2000). A group-level model of safety climate: Testing the effect of group climate on microaccidents in manufacturing jobs. *Journal of Applied Psychology*, 85(4), 587–596.