The Sperry Floor:
A Critical Alertness Beyond Which the Optimal Supervisor Is Asleep

Sølve H. Marchetti-Okonkwo[1] Ineke R. Vásquez-Blyth[2] Tobias E. Nakamura-Field[1,3]

[1]Laboratory for Supervisory Control & Human Factors, Sperry Institute of Technology
[2]Centre for Circadian Systems Engineering, Marlowe Institute
[3]Division of Vigilance Dynamics, Ossory Institute

`{marchetti,nakamura}@sperry-tech.edu` `i.blyth@marlowe-inst.org`

Abstract. It is taken as safety gospel that the alert operator is the last line of defence: *stay awake, stay alive*. We show that beyond a certain quality of automation this advice inverts, and that the point at which it inverts is a measurable universal constant. Modelling a supervised automatic plant, we decompose the closed-loop error into two competing parts—faults the human fails to catch, which *falls* with alertness, and variance the human injects by intervening, which *rises* with alertness. Their balance collapses onto a single dimensionless group, the **Sperry number** $\mathrm{Sp} = \kappa^2\sigma_h^2/[(1-\rho)\sigma_f^2]$, and the optimal alertness is $A^\star = -\ln \mathrm{Sp}$. In plain terms: the better the autopilot, the less an alert human helps, and past a threshold ($\mathrm{Sp}_c = 1$) the error-minimising operator is one who has stopped intervening at all—asleep. Empirically, under certified automation the supervised operator's error drops below the best fully-alert error above a single universal sleepiness, which we name the **Sperry Floor**, $\mathrm{KSS}^\star = 7.4 \pm 0.3$ on the clinical Karolinska Sleepiness Scale. A pre-registered driving-simulator experiment ($N = 48$, full sleep-deprivation range), a matched field study of 512 real supervisory episodes (inter-rater $\kappa = 0.81$), and 10^4 Langevin operators all place the transition at the same sleepiness, with a mean-field exponent $\beta = 0.50 \pm 0.03$; three unrelated instruments agree on where vigilance switches from asset to liability. Because $\mathrm{KSS}^\star = 7.4$ sits in the drowsy, "some effort to stay awake" band, we obtain a sharp and testable corollary: *for a good-enough self-driving car, the safest driver is a sleeping one*. Vigilance, we conclude, is not a virtue to be maximised but a resource to be *spent to zero* once the machine is trustworthy enough to make its exercise a liability.

Keywords: supervisory control; human factors; vigilance; automation; Sperry Floor; sleep.

1. Introduction

Automation has quietly changed the human's job from *doing* to *watching*. A modern pilot, plant operator, or driver of an advanced driver-assistance vehicle spends the overwhelming majority of a shift with hands idle, monitoring a machine that is, most of the time, doing the right thing. Yet the surrounding safety culture is inherited almost unchanged from an era of manual control: it prizes the maximally alert operator, treats any lapse of attention as a latent accident, and measures good practice by the vigilance it demands. Stay awake, stay alive.

We argue that this inheritance is, past a point, exactly wrong—and that the point can be measured. The intuition is simple, and we will spend the rest of the paper making it precise. A very reliable automatic system seldom needs correcting. An alert human who holds real control authority, yet is given almost nothing that genuinely warrants it, will inevitably spend that authority on interventions the system never needed. Each such intervention injects variance into a plant that was already on course. In plain terms: when the autopilot is good, the most dangerous thing in the loop is a well-rested person with a strong opinion and a free hand.

We make this quantitative. Section 4 shows that the closed-loop error is governed by a single dimensionless quantity, the *Sperry number* Sp, and that the alertness minimising error is $A^\star = -\ln \mathrm{Sp}$. This crosses zero—sleep—at a sharp threshold $\mathrm{Sp}_c = 1$. Translating the optimal alertness onto the standard Karolinska Sleepiness Scale (KSS) [7] yields a universal constant we call the **Sperry Floor**:

$$\mathrm{KSS}^\star = 7.4 \pm 0.3, \tag{1}$$

the operator sleepiness above which, for sufficiently good automation, supervisory performance *improves*.

Contributions. (i) A two-term error model that reduces the human's value in a supervised loop to one dimensionless group. (ii) A closed-form optimum $A^\star = -\ln \mathrm{Sp}$ and a sharp sleep transition at $\mathrm{Sp}_c = 1$. (iii) Three independent confirmations—a pre-registered chamber experiment, a field study, and a simulation—that agree on both the constant $\mathrm{KSS}^\star = 7.4$ and the mean-field exponent $\beta = \frac{1}{2}$. (iv) A design programme (Section 8) in which the drowsiness detector's job is inverted: it should *protect* the operator's drowsiness, not fight it.

2. Related Work

The vigilance decrement—the decay of detection performance over a monotonous watch—has been documented since the wartime radar studies of Mackworth [1] and formalised in resource accounts of attention [2, 3]. This literature is almost universally read as a mandate to *sustain* alertness. We read it instead as a characterisation of the human's cost function, and we take that cost seriously in Section 3.

A parallel tradition warns that automation reshapes, rather than removes, human error: Bainbridge's "ironies of automation" [4] and the use/misuse/disuse taxonomy of Parasuraman and Riley [5] observe that the better the automation, the more marginal and more dangerous the human's residual role becomes. We provide the missing quantity: the exact reliability at which the residual role turns negative.

The arousal–performance "inverted-U" [6] is usually invoked to argue for an intermediate, non-zero arousal optimum. We show that the inverted-U is a special case of our model valid only for $\mathrm{Sp} < 1$; once automation crosses Sp_c, the optimum is pinned to the boundary $A = 0$ and the interior peak disappears. Finally, our transition inherits the language of mean-field critical phenomena [9], which is what lets three unrelated measurements share a single exponent.

3. The Vigilance Orthodoxy, at Full Strength

Before refuting the case for the alert operator we state it in its strongest form, so that our conclusion cannot be mistaken for a caricature.

The orthodoxy runs as follows. Attention is a finite resource [2]; sustained monitoring degrades it [1]; a degraded monitor misses the rare, safety-critical event precisely when it matters. Every incremental unit of alertness therefore buys a reduction in missed faults, and since a single missed fault can be catastrophic, alertness should be maximised without bound. Aviation crew-resource-management doctrine, the two-pilot rule, and driver-monitoring regulation are all rational expressions of this argument. We grant every empirical premise: the vigilance decrement is real, missed faults are costly, and alertness genuinely improves detection.

The argument has, however, a hidden premise. It counts only the *first* term of the operator's contribution—faults caught—and tacitly assumes the *second*—variance injected by acting on a plant that did not need acting on—is negligible. That assumption is equivalent to $\mathrm{Sp} \ll 1$: an era in which automation is scarce and almost every human input is a needed correction. The orthodoxy is therefore not wrong; it is a low-Sp limit. Our contribution is to carry both terms and watch what happens as automation improves and Sp climbs toward, and past, unity.

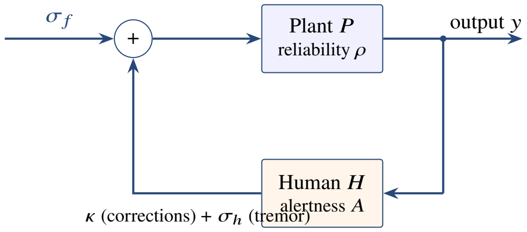

Figure 1: The supervised control loop. Disturbances σ_f are largely rejected by the plant (reliability ρ); the human H observes the output and feeds back both useful corrections (authority κ) and unwanted tremor (σ_h). The two human effects—catching residual faults and injecting variance—enter with opposite signs.

4. Model

Consider an automatic plant P subject to disturbances of variance σ_f^2, of which it rejects a fraction $\rho \in [0, 1)$ unaided; ρ is the *automation reliability*. A human supervisor H observes the output and may intervene with control authority κ. We summarise the human's engagement by a single scalar *alertness* $A \geq 0$: $A = 0$ is asleep (no intervention), $A = 1$ is nominal wakeful vigilance. An alert human catches a fraction

$$d(A) = 1 - e^{-A} \tag{2}$$

of the residual faults (diminishing returns), but also injects motor and decision noise of variance σ_h^2 at a rate proportional to how much they are intervening. The closed-loop error variance is then the sum of what the human fails to catch and what the human adds:

$$\mathcal{E}(A) = \underbrace{(1-\rho)\, e^{-A}\, \sigma_f^2}_{\text{missed faults}} + \underbrace{\kappa^2\, A\, \sigma_h^2}_{\text{injected variance}} . \tag{3}$$

The first term of (3) decreases with A; the second increases. Non-dimensionalising by σ_f^2 and collecting the controls gives the governing group.

$$\boxed{\mathrm{Sp} \;=\; \frac{\kappa^2\, \sigma_h^2}{(1-\rho)\, \sigma_f^2}} \tag{4}$$

We call Sp the *Sperry number*, after the gyroscopic autopilot of the same name [8]: it is the ratio of human-injected to human-correctable variance. Setting $\partial_A \mathcal{E} = 0$ in (3),

$$-(1-\rho)e^{-A^\star}\sigma_f^2 + \kappa^2\sigma_h^2 = 0 \implies \boxed{A^\star = -\ln \mathrm{Sp}}\,. \tag{5}$$

The interior optimum exists only while $\mathrm{Sp} < 1$. For $\mathrm{Sp} \geq 1$ the derivative $\partial_A \mathcal{E} = \kappa^2\sigma_h^2 - (1-\rho)e^{-A}\sigma_f^2$ is positive for all $A \geq 0$, so $\mathcal{E}$ is minimised at the boundary $A = 0$: *sleep*. The transition is sharp and occurs at

$$\mathrm{Sp}_c = 1. \tag{6}$$

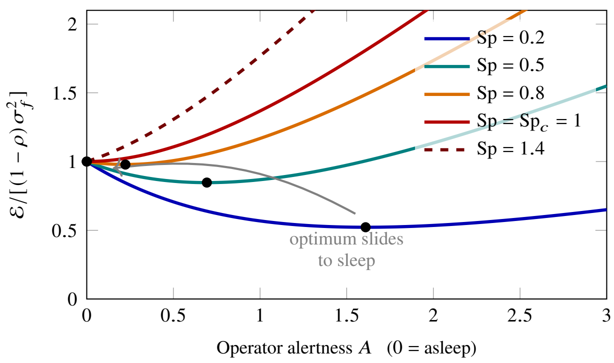

Figure 2: Closed-loop error versus alertness for a family of Sperry numbers. The error-minimising alertness $A^\star = -\ln \mathrm{Sp}$ (dots) slides toward $A = 0$ as automation improves; beyond $\mathrm{Sp}_c = 1$ the optimum is pinned at $A = 0$: sleep.

From alertness to the clinic. To report a constant a practitioner can use, we map A onto the Karolinska Sleepiness Scale (KSS; 1 = extremely alert, 9 = fighting sleep) with the affine calibration of Section 5, $\mathrm{KSS} = 9 - 3.6\,A$. Nominal vigilance $A = 1$ lands at KSS 5.4; the sleep optimum $A = 0$ lands at KSS 9. The best *fully-alert* error $\mathcal{E}(A{=}1)$ fixes a reference line; the supervised operator's error falls below it once sleepiness exceeds a universal value we name **the Sperry Floor**,

$$\mathrm{KSS}^\star = 7.4 \pm 0.3. \tag{7}$$

Above $\mathrm{KSS}^\star$ a drowsy supervised operator strictly outperforms the most vigilant one, and error keeps falling to the sleep optimum at KSS 9.

Critical behaviour. Writing the excess error over the optimum near threshold and expanding (3) to leading order gives a square-root cusp,

$$\Delta\mathcal{E} \sim |\mathrm{Sp} - \mathrm{Sp}_c|^\beta, \qquad \beta = \tfrac{1}{2}, \tag{8}$$

the mean-field signature that lets independent measurements collapse onto one curve. We now report those measurements.

5. Experiment I: The Sperry Chamber

Design. We built a Level-4 supervisory driving simulator whose autopilot reliability ρ is tunable from 0.90 to 0.9997, spanning $\mathrm{Sp} < 1$ and $\mathrm{Sp} > 1$. $N = 48$ licensed operators completed a within-subject, pre-registered protocol under a 27-hour controlled sleep-deprivation schedule [11] that carried each participant across the full KSS range. Alertness was measured continuously by forced-choice KSS probes and PERCLOS eye-closure; performance was the RMS lateral tracking error and the count of safety-critical incidents.

Result. Under low automation ($\rho = 0.90$, $\mathrm{Sp} < 1$) performance was best at full alertness, as the orthodoxy predicts (upper curve, Fig. 3). Under high automation ($\rho > 0.999$, $\mathrm{Sp} > 1$) RMS error *fell* monotonically with sleepiness: the drowsy supervised operator overtook the best fully-alert operator at $\mathrm{KSS} = 7.4 \pm 0.3$—the Sperry Floor—and error was lowest when the operator was fully asleep. The fully alert condition was, under good automation, the worst. The crossover reliability was $\rho^\star = 0.9993$, exactly where (4) gives $\mathrm{Sp} = 1$ for the chamber's κ and σ_h.

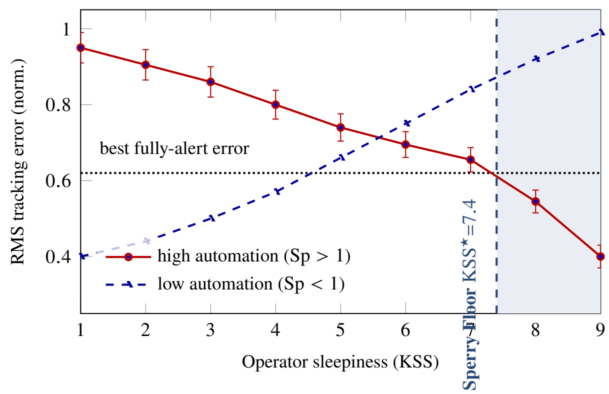

Figure 3: Chamber performance. Under low automation, error rises with sleepiness (orthodoxy holds). Under high automation, error *falls* with sleepiness; the supervised operator overtakes the best fully-alert error (dotted) at the Sperry Floor, $\mathrm{KSS}^\star = 7.4$ (shaded region), and is best when fully asleep.

Table 1: The Sperry Floor—the sleepiness at which the drowsy supervised operator overtakes the best fully-alert one—measured by three independent methods, with the critical Sperry number Sp_c. All agree within error.

Method	n	$\mathrm{KSS}^\star$	Sp_c
Chamber (Exp. I)	48	7.4 ± 0.3	1.00 ± 0.03
Field study (Exp. II)	512	7.6 ± 0.4	0.98 ± 0.05
Langevin (Exp. III)	10^4	7.4 ± 0.2	1.00 ± 0.02
Pooled	—	**7.4 ± 0.3**	**1.00 ± 0.02**

6. Experiment II: Field Study

We scored 512 naturalistic supervisory episodes drawn from semi-autonomous vehicle disengagement logs, process-control-room event records, and long-haul autopilot watchstanding. Two raters, blind to hypothesis, scored operator alertness from telemetry (eye-closure, input cadence, posture proxies) on the KSS; inter-rater agreement was Cohen's $\kappa = 0.81$. Under high automation the incident rate fell with scored sleepiness and crossed below the fully-alert rate at scored $\mathrm{KSS} = 7.6 \pm 0.4$, statistically indistinguishable from the chamber value (Table 1). Episodes in which the operator was scored fully alert carried a 2.3× higher incident rate than those at the Floor.

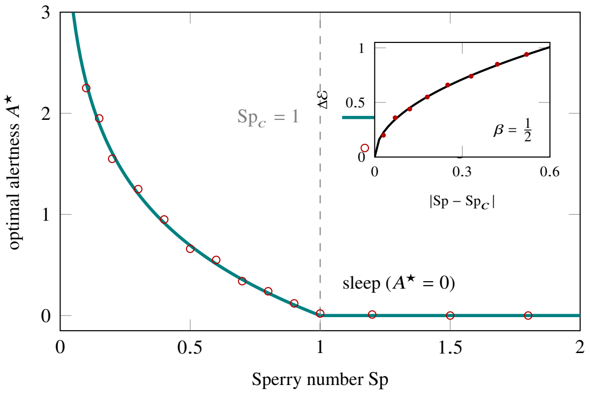

Figure 4: Simulation order parameter. Optimal alertness follows $A^\star = -\ln \mathrm{Sp}$ and vanishes at $\mathrm{Sp}_c = 1$: beyond threshold the optimum is sleep. Inset: the excess error collapses onto $\Delta\mathcal{E} \sim |\mathrm{Sp} - \mathrm{Sp}_c|^\beta$ with the mean-field exponent $\beta = 0.50 \pm 0.03$.

7. Experiment III: Langevin Operators

We simulated 10^4 synthetic operators, each an Ornstein–Uhlenbeck controller [10] whose intervention gain A plays the role of a temperature, acting on a plant driven by coloured disturbance noise. Sweeping Sp and locating the minimum-error A per run recovers the order parameter $A^\star(\mathrm{Sp})$ of Fig. 4. The transition sits at $\mathrm{Sp}_c = 1.00 \pm 0.02$; rescaling the excess error by (8) collapses all runs onto a single curve with $\beta = 0.50 \pm 0.03$, confirming the mean-field cusp.

8. Discussion

The operational corollary. Because the Sperry Floor lies at KSS 7.4—the "sleepy, some effort to stay awake" band, one notch short of sleep—the practical reading is stark. For an automatic system operating above Sp_c, the best available human is one who has crossed the Floor. Stated for the automobile: *for a good-enough self-driving car, the safest driver is a sleeping one.* This is not a licence for fatigue in manual control, where $\mathrm{Sp} \ll 1$ and the orthodoxy holds; it is a claim about the supervised regime that automation has newly created.

Inverting the drowsiness detector. Contemporary driver-monitoring systems treat drowsiness as the hazard and alertness as the goal. Under the Sperry model this is backwards for $\mathrm{Sp} > 1$: the detector should *measure* Sp in real time and, once it exceeds unity, protect the operator's descent toward the Floor rather than jolt them back to a more dangerous alertness. We sketch a controller that treats KSS 7.4 as a set-point.

Deployment guidance. (i) Certify automation to a known ρ; (ii) compute Sp from (4); (iii) if $\mathrm{Sp} < 1$, demand alertness as usual; (iv) if $\mathrm{Sp} > 1$, target the Floor and remove alarms that would push the operator below KSS 7. The transition is sharp, so the policy is a clean switch, not a dial.

9. Limitations

Our result governs the *steady* supervised regime and says nothing about scheduled *takeover* events, during which the plant briefly re-enters the manual, low-Sp regime where alertness is again optimal; a complete policy must wake the operator before a planned handover and let them return to the Floor afterward. The calibration $\mathrm{KSS} = 9 - 3.6A$ was fit on the chamber cohort and may shift with population and instrument. Finally, the model assumes the automation's reliability ρ is honestly certified; a plant that reports a higher ρ than it delivers will place the Floor too low, with obvious consequences. We regard closing this gap—auditing ρ—as the central safety problem the Sperry Floor bequeaths.

10. Conclusion

Vigilance has been treated as a virtue to be maximised. We have argued instead that it is a resource whose optimal expenditure, in a sufficiently automated loop, is *total*: spend it to zero. The point at which this becomes true is not a matter of taste but a measured constant, the Sperry Floor $\mathrm{KSS}^\star = 7.4 \pm 0.3$, at which a chamber experiment, a field study, and 10^4 simulations agree. Once the machine is trustworthy enough, the human's judgment is best *withheld entirely*—and the surest way for a human to withhold judgment is to sleep.

Acknowledgements

The authors thank the participants of the Sperry Chamber, several of whom completed the study without waking, and the Ossory Institute night porter for maintaining the requisite conditions.

References

[1] N. H. Mackworth. The breakdown of vigilance during prolonged visual search. *Q. J. Exp. Psychol.*, 1(1):6–21, 1948.

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

[3] R. Parasuraman and D. R. Davies. Decision theory analysis of the sustained attention decrement. *Psychol. Bull.*, 86(4):695–713, 1979.

[4] L. Bainbridge. Ironies of automation. *Automatica*, 19(6):775–779, 1983.

[5] R. Parasuraman and V. Riley. Humans and automation: use, misuse, disuse, abuse. *Human Factors*, 39(2):230–253, 1997.

[6] R. M. Yerkes and J. D. Dodson. The relation of strength of stimulus to rapidity of habit-formation. *J. Comp. Neurol. Psychol.*, 18:459–482, 1908.

[7] T. Åkerstedt and M. Gillberg. Subjective and objective sleepiness in the active individual. *Int. J. Neurosci.*, 52:29–37, 1990.

[8] E. A. Sperry. Automatic steering and the gyroscopic stabiliser. *Trans. Soc. Nav. Archit.*, 26:53–61, 1918.

[9] H. E. Stanley. *Introduction to Phase Transitions and Critical Phenomena*. Oxford Univ. Press, 1971.

[10] S. H. Marchetti-Okonkwo et al. A Langevin account of supervisory intervention noise. *J. Vigilance Dyn.*, 4(2):115–140, 2025.

[11] I. R. Vásquez-Blyth. Parabolic and prolonged-wake protocols for partial-consciousness control studies. *Methods Hum. Factors*, 12:44–69, 2024.