NOCICEPTION AS A CONTROL PRIMITIVE: AFFERENT CHANNELS AND NOCICEPTIVE MEMORY FOR AGENTS DEPLOYED IN ONE BODY

Anonymous authors
Paper under double-blind review

ABSTRACT

An agent deployed in a single body cannot learn how fast that body wears, because every trial that would reveal its wear resistance wears the body it would protect. We study this *epoch-one* setting, in which the parameters of a fixed-weight policy are set before the body is drawn and never updated in life. The agent carries a load-gated nociceptive channel and a memory that retains what was felt. We prove that felt cost moves the allocation to the best-*paid* work not yet felt rather than the gentlest, that an agent without retention never sees the felt-cost constraint bind, and that the channel pays only where the threat is individually unpredictable, cheap to avoid and expensive to ignore. We measure per body, setting the agent with channel and memory against the same individual without them, where neither carries a schedule learned across lives. On $2{,}000$ simulated floor-layer knees, with wear anchored to published loss rates, feeling, retaining and substituting extends the working life from age 55.2 to 59.6 and raises career output from 33.7 to 36.1. 69.3% of bodies gain and **none lose**. A body that feels but retains nothing past the day gains one of the $+4.4$ years, and retention carries the rest. A population-trained agent gains $+0.65$ years from the same channel at -0.54 output. The difference is what a species prior already supplies, and a single body has none. The two are related by an identity, the ablation mean reporting $(1 - \chi)$ of the per-body value with χ the share a blind schedule already captures, so we report both. Where the regime map predicts value, a care robot sextuples its certified service life and a field-anchored fleet writes off 0.15 of its machines instead of 0.55. Where it predicts none, a rover gains little over blind caution, so the map holds in both directions.

1 INTRODUCTION

Gradient-based optimization needs its objective sampled many times (Schulman et al., 2017), whereas an individual has one life. The events that dominate a lifetime objective occur less than once within it (Lipton et al., 2016), and the trials that would identify its own wear resistance are the ones it cannot afford, because each wears the body it would protect, an irreversibility that safe exploration assumes known or learnable (García & Fernández, 2015; Turchetta et al., 2016; Grinsztajn et al., 2021; Turner et al., 2020; Khetarpal et al., 2022). Individuals are nevertheless competent, and this paper asks what an agent must carry into one life for that to be possible and measures what the answer is worth to the body that carries it.

Two timescales have to be kept apart. Optimization belongs to the slow loop, phylogeny in biology and population training in our experiments, and adaptation to the fast loop, one deployment life. A policy trained over a distribution of bodies and environments (Tobin et al., 2017; Duan et al., 2016; Rakelly et al., 2019) installs a prior over bodies and never observes the draw that is realized. In the fast loop, adaptation cannot be optimization, because with fewer than one informative event per life the hidden wear resistance is not recoverable from safe experience (Prop. 2), so the premise of Bayes-adaptive control, a latent observed often enough to estimate (Duff, 2002; Ghavamzadeh et al., 2015; Kumar et al., 2021), fails by construction. What remains is memorization, conditioning on what has been felt rather than on an estimate. We call this regime **epoch one**: one deployment life, parameters fixed before the embodiment is known, no in-life updates, the setting of single-life RL

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

(Chen et al., 2022) and of in-context adaptation with fixed weights (Laskin et al., 2023; Lee et al., 2023; Bauer et al., 2023).

We treat nociception as a control primitive, as homeostatic RL and artificial pain do (Keramati & Gutkin, 2014; Man & Damasio, 2019; Kuehn & Haddadin, 2017), but read differently. It is a coarse readout of tissue state that is *load-gated*. The signal rises with the activity that produces it, so it identifies its own cause and affords *substitution* within the repertoire a trade or a mission permits, rather than reduced effort. Because what is felt can be retained, as a conditioned aversion is (Bouton, 2004; Hasenbring & Verbunt, 2010), one event can redirect the allocation for the rest of the life. What is under test is therefore a loop closed through the body. Pain moves the allocation, the allocation sets the load, the load sets both how fast the tissue wears and how strong the next signal is, and a signal that fades with the work that caused it leads the body to resume that work. The four links, **pain, memory, substitution, and a longer working life**, are traversed many times in one life, and memory is what keeps the loop from cycling. Pain without retention moves a body to the best-*paid* work it can still do and returns it there whenever the signal dips, the same optimism that drives exploration for rewards (Auer et al., 2002; Brafman & Tennenholtz, 2002) acting on a cost it has not yet felt. What is retained moves it to the gentlest and holds it there (Lemma 1, and §5.2 counts the returns). The loop is conditional at every link and operates only where the repertoire admits a gentler allocation that still clears the output floor (Thm. 1, Prop. 3).

The clinical counterpart is Charcot arthropathy, in which an insensate patient reloads an injured joint where one with intact sensation protects it for decades, on the same species prior (Nagasako et al., 2003; Rogers et al., 2011; Woolf, 2010). We measure per agent, setting the apparatus against the same individual without it, where neither arm carries a schedule learned across lives. On a floor layer's knee, working at full effort reaches age 55.2 and 33.7 in career output. Sparing from the first day reaches age 61.5 and output 34.4. Feeling, retaining and substituting reaches age 59.6 and output 36.1 (Table 1). Over $2{,}000$ paired bodies 69.3% gain years and **none lose any**. The body that feels but retains nothing past the day gains one year, and retention carries the rest. A care robot, retired when its certification lapses, serves 323 rather than 51 months when it reads the signal, and there retention subtracts 34 of them (Prop. 12).

What must a fixed-weight agent carry through one life in a body that wears, and what is that worth to the body? Three results answer this. *Content of the state.* Under an optimistic prior a felt-cost constraint sends the agent to the best-paid untried option, and without retention the constraint never binds, since the signal exists only while the harmful work is done (Lemma 1, Cor. 3). Retention is therefore the precondition of sensing, and the state stores a threat, kept after the activity is set down. *Regime.* The channel pays only where the threat is individually unpredictable, cheap to avoid and expensive to ignore, and there the optimum is protective if and only if sensing, persistence and tail-sensitivity hold jointly (Prop. 3, Thm. 1). All three conditions can be checked before a world is run, and a world built after the prediction confirmed it. *Measurement.* The value lies on the bodies at risk, so a population mean reports q times what an at-risk body experiences, $q = 0.693$ here (Rem. 1), and an ablation of a population-trained agent reports $(1 - \chi)$ of the per-body value, χ the share a blind schedule already captures (Prop. 1). χ is a property of the world, since in a minimal MDP with no body and no learner the channel is worth the same to every individual while the ablation contrast moves by a factor of twelve, so we measure per body, against the same body without in-life state. The paper thus names a setting, shows that in it retention is a precondition of sensing rather than a refinement, gives conditions for when a body model pays that are decidable before the world exists, and supplies a measurement identity that applies beyond its subject.

2 Related work

Four bodies of work touch the setting, and each rests on an assumption that epoch one removes. *Adaptation within a life.* Episodic control stores returns of visited states (Blundell et al., 2016; Pritzel et al., 2017), meta-RL and rapid motor adaptation infer a latent task or embodiment from repeated episodes or recent history (Duan et al., 2016; Rakelly et al., 2019; Kumar et al., 2021), single-life RL adapts by transfer from the same task distribution (Chen et al., 2022), and in-context adaptation with fixed weights adapts through state that grows within the episode (Laskin et al., 2023; Lee et al., 2023; Bauer et al., 2023). All of them assume that the latent which matters can be estimated from the experience the life affords, or that similar lives exist to transfer from. Here

the latent is a wear resistance that safe experience cannot identify (Prop. 2), so the within-life state cannot be an estimate. We keep the policy class of the last line, fixed weights and a growing state, and change what the state stores to a felt cost per activity, retained after the activity is set down.

Safety and constraints. Risk-sensitive and constrained RL change the objective (Tamar et al., 2015; Chow et al., 2015; Altman, 1999; Achiam et al., 2017; García & Fernández, 2015), safe exploration learns an unknown constraint by approaching it, pairing optimism on reward with pessimism on untried cost (Turchetta et al., 2016; Wachi & Sui, 2020; Efroni et al., 2020; Liu et al., 2021; Bura et al., 2022), conservative bandits hold a baseline while they explore (Wu et al., 2016; Kazerouni et al., 2017), and shields or barrier functions enforce a constraint given in advance (Alshiekh et al., 2018; Ames et al., 2019). Our constraint is neither given nor modelled but felt, and felt only while the harmful work is done. That is why the optimism these methods rely on is the hazard here, since to a body that retains nothing an untried option looks free (Lemma 1). Pessimism on untried cost is the fixed-endowment row of Table 1, and the risk-sensitive objective runs as a CVaR arm in the fleet (App. C.1).

Engineering under degradation. Condition-based maintenance estimates remaining useful life from monitored condition (Jardine et al., 2006; Si et al., 2011; Saxena et al., 2008), and that estimate drives inspection and repair as a POMDP (Andriotis & Papakonstantinou, 2019) or service as a restless bandit (Whittle, 1988). The degradation model behind such an estimate is fitted on a population of units, which is our slow loop, and the population-trained blind arm of Table 1 is a schedule of this kind. Our question is what one unit gains beyond it when its wear rate is drawn per unit and never observed (Prop. 1).

Signals from the body. Homeostatic and interoceptive RL regulate toward a physiological setpoint (Keramati & Gutkin, 2014; Man & Damasio, 2019; Craig, 2002), artificial pain supplies a nociceptive channel to a robot (Kuehn & Haddadin, 2017), intrinsic fear penalizes states that precede catastrophe (Lipton et al., 2016), and damage-recovery repertoires search for a gait after an injury (Cully et al., 2015). These are the closest in spirit, but they use the signal as a reward term or as a trigger for search, whereas we use it as a load-gated constraint with retention and ask what it is worth to one body and in which worlds. The formal tools are approximate information states, which say what a compressed state must retain (Åström, 1965; Subramanian et al., 2022), partial monitoring (Bartók et al., 2014), high-confidence off-policy evaluation (Thomas et al., 2015), and the stochastic-programming separation of the value of the stochastic solution from perfect information, the identity behind the schedule confound (Madansky, 1960; Birge & Louveaux, 2011; Duff, 2002; Ghavamzadeh et al., 2015).

3 THE SINGLE-LIFE MODEL

The model has three parts (Fig. 1). The fast loop is one life, a world with a hidden body, an afferent layer that reads it and a decision level that allocates the work. The slow loop sets the parameters of the last two before the life begins and receives nothing back. The experimenter stands outside both loops and compares the same body with and without the channel. We define the world of the fast loop first, then the policy class the slow loop sets, then the two values the experimenter reads, then the agent.

The fast loop, its world. Let $\{\mathcal{M}(\theta, e)\}$ be a family of POMDPs indexed by embodiment θ and environment e, with training distribution $p(\theta, e)$. The embodiment θ is the body the agent is dealt and cannot change. The environment e is what the world demands of that body: the task mix it must keep, the floors and caps, and the event statistics. The load the body carries is neither, it is chosen inside the life. The setting rests on the hiddenness of the realized (θ, e), which is unidentifiable from safe experience (Prop. 2), and requires no genomic assumption. Each life has horizon H, an irreversible integrity state M_t and a repertoire A of activities from which each period a portfolio is composed. Activity j has load ℓ_j on the site and yield y_j to the work, and the two do not rank alike, since the best-paid day type wears the joint. Each period a life selects shares $a_t \in \Delta(A)$ under two constraints, so substitution is bounded from both sides. The mix must deliver at least $\underline{y}$ or the work is lost, and wear also removes options, since activities whose demands exceed current capability leave A. Integrity falls with the work done:

$$M_{t+1} \;=\; M_t \;-\; \frac{d(a_t, M_t)}{\kappa(\theta)}. \tag{1}$$

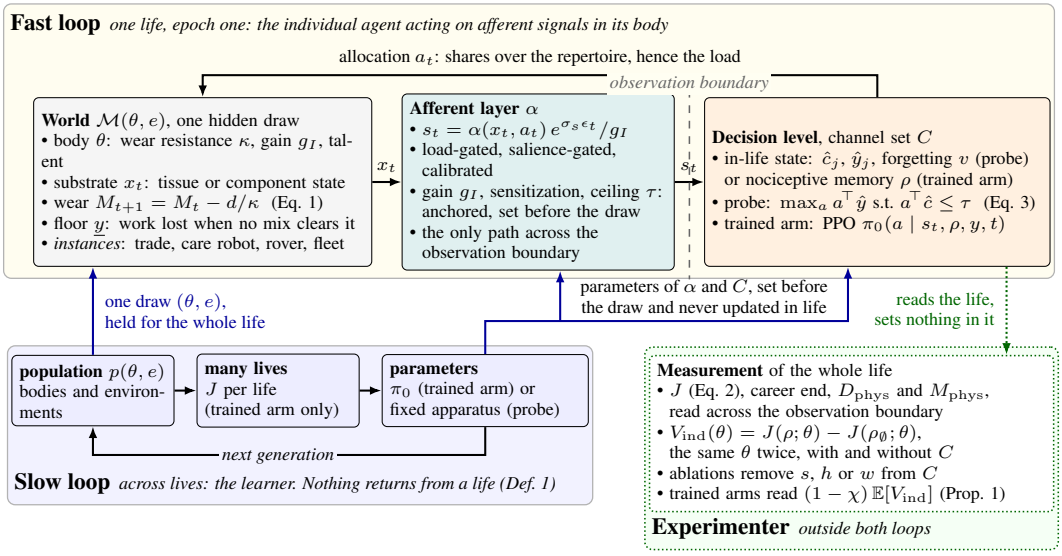


Figure 1: **Architecture.** The fast loop is one life inside a single hidden draw of the world. The dashed line is the observation boundary, crossed only by the afferent signal. The decision level allocates the day, which sets the load, the wear and the strength of the next signal, and its in-life state is the recorded felt cost per activity (probe) or a nociceptive memory (trained arm). The slow loop feeds the fast loop only before the life begins, with the draw (θ, e) and with the parameters of both the afferent layer and the decision level, and receives nothing back (Def. 1). The experimenter reads the whole life across the observation boundary and sets nothing inside it.

$d(a_t, M_t)$ is the wear of ordinary use, irreversible and deterministic in the activity and the state already reached. In the wear worlds it is the portfolio's hazard-weighted load times an amplification that rises as the substrate degrades, net of a slow repair term and with no acceleration in the damage state itself, and its rate is anchored to published loss rates of the substrate (App. E, Maschek et al., 2014; Roth et al., 2017). The drawn wear resistance $\kappa(\theta)$ divides it, so one body wears faster than another under identical work, and κ is what a life would need to know about itself and cannot (Prop. 2). Some worlds additionally carry acute events, an extension that is stated in App. B and run by the fleet of §5.1.

A career ends in the first period in which no admissible mix clears $\underline{y}$, which under Eq. 1 is the analogue of an event with probability near one rather than near zero. The chain has four links. The signal identifies the activity that caused it. A trace retains that identity once the activity is set down and the signal with it. The repertoire permits the activity to be traded, and the floor decides for how long. Three are properties of the world, ℓ_j, y_j and $\underline{y}$, and the trace is the one property of the agent, entering the decision as the forgetting rate v of Eq. 3 (Fig. 5 in App. E).

The slow loop. Def. 1 fixes what the slow loop may set, and with the reference arm also what the experimenter may compare.

Definition 1 (Epoch-one policy class)**.** *At deployment one hidden pair (θ, e) is drawn and held for the whole life. $\Pi(C)$ contains policies measurable with respect to the in-life filtration $\mathcal{F}_t^C$ of channel set C, with fixed parameters, hence no dependence on the realized θ and no updates. The reference arm $\rho_\emptyset$ is the member with $C = \emptyset$ and no schedule chosen with knowledge of which work is gentlest.*

The slow loop supplies the prior π_0, and deployment competence is what C adds. An individual has no knowledge of which work is gentlest, so a constant sparing policy is inadmissible as $\rho_\emptyset$, because comparing against one measures that knowledge instead.

The experimenter. Two values follow from the definition, and they must not be confused. $V(C \mid C')$, the gain of the best policy reading C over the best reading C', measures information over the species prior and is nil wherever π_0 can carry the schedule. $V_{\text{ind}}(\rho; \theta) := J(\rho; \theta) - J(\rho_\emptyset; \theta)$ is the paired per-body gain on identical draws, reported as a distribution. What an ablation ablates is not the naive $\rho_\emptyset$ but, trained over the population, the best *schedule*. Writing χ for the ratio of what that schedule

captures to the individual gain, the two designs differ by exactly that term (Prop. 1). We call the gap the **schedule confound**.

The fast loop, its agent. One layer, the *afferent layer*, separates the biology from everything above it. The substrate state x_t is domain-specific and never seen by the decision level. The afferent map $\alpha : (x_t, a_t) \mapsto s_t \in [0,1]^k$, one entry per activity, transduces it onto one common scale, $\alpha(x,a) = \ell(a)\, S(x)/\bar{\alpha}$ clipped per activity, with $\bar{\alpha}$ a population constant. Three properties follow. *Load-gating*: damage is felt when the activity is worked and falls when work substitutes away from it, which makes the signal actionable. *Calibration*: one scale serves policy input, memory write and cost. *Salience-gating*: events saturate the scale, so the spike alone triggers a write. The map is a product, so the same action is felt more as the body wears and the constraint tightens on its own as the reserve falls, without the life computing how much of itself remains. *In this sense nociception is a control primitive*, since the body registers the depletion of a resource on the scale of the decision that depletes it.

During a deployment the decision level observes its own body and history only through a channel set C, holding the felt signal s_t and a nociceptive memory over activities built from it. The dashed line of Fig. 1 between the two is the *observation boundary*, and what lies left of it is never observed.

$$s_t = \frac{\alpha(x_t, a_t)}{g_I}\, e^{\sigma_s \epsilon_t}, \qquad J = \textstyle\sum_t y(a_t) - w \sum_t s_t. \tag{2}$$

Here $y(a) = \sum_j a_j\, y_j$ is the yield of the mix as the body can still deliver it, each y_j reduced by what wear has taken from the capability the activity needs, and $\alpha(x_t, a_t)$ is the afferent map of Fig. 1. The individual gain g_I and the wear resistance $\kappa(\theta)$ of Eq. 1 are drawn separately, both log-normal. The spread of the gain and the reading noise σ_s are anchored per world on the between-person and the within-person variability of the felt signal, while the spread of κ is set rather than anchored and is therefore driven as an axis (App. E). No substrate variable enters the observation, so the afferent map is the only path across the observation boundary of Fig. 1 and the observation does not recover κ (Prop. 2). The objective trades realized earnings against the accumulated felt signal, weighted by the valuation trait w, so $w{=}0$ recovers pure output-maximization. Here y is what the agent collects and not what it produces, because losing the work does not end the stream. An outside option follows, re-employment for a worker, a residual use for a machine, nothing for a written-off unit, and its level prices the loss of the work (App. E). The memory is deliberately almost nothing, a conditioned aversive trace with extinction and recovery, recording where and how much but never what or when, and the only state that grows within a life (App. B.1), and nothing affective is implemented.

Eq. 2 prices the signal in the objective, which is how every arm is *scored*. It is not how the signal reaches the allocation. The decision level carries two numbers per activity, an estimate $\hat{y}_j$ of what it still delivers and an estimate $\hat{c}_j$ of what working it costs to feel, and each period it solves

$$\max_{a \in \Delta(A)} \sum_j a_j\, \hat{y}_j \qquad \text{subject to} \qquad \sum_j a_j\, \hat{c}_j \;\le\; \tau, \tag{3}$$

under the availability limits the trade puts on each group of activities and the floor $\underline{y}$ that decides whether the work is kept at all. Both estimates are written by working and by nothing else. At the end of a period they take the values just delivered and just felt on every j the period worked, while every other j relaxes toward the endowment $(y_j, \hat{c}_j^0)$ at a forgetting rate $v \in [0,1]$. That rate is the second link of the chain: $v = 1$ forgets everything but what was worked last, $v = 0$ forgets nothing, and these are the last two arms of Table 1. The first feels the work it is doing and retains nothing past the day, where the day is the allocation of one period, so that the next allocation is chosen on what the last day felt and on nothing older, whereas the second retains what it felt. The ceiling τ is the felt load of a fresh body at its observed task mix, scaled by the ratio of the stress at which the substrate begins to fail to its intact peak stress, both taken from the literature of the world (App. E), giving $\tau = 0.38$ in the trade.

Provenance of the ceiling. A price with $w > 0$ and no constraint could carry the same information. We run the ceiling because τ is a threshold in contact mechanics whereas a price would have to be chosen (Prop. 8 states what choosing it wrongly costs), and Lemma 1 and Cor. 3 are stated over Eq. 3. The trade's agent carries $w = 0$, so its output is what it earned, while J still scores every arm in every world.

4 THEORY

Two results carry the chronic world and chain, Lemma 1, which says where a cost-constrained agent escapes to and why zero retention leaves the constraint vacuous, and Prop. 4, which prices sensing net of the wear spent acquiring it.

Write Δ_b for the working life a protective response can add to one body and q for the share of bodies whose working life is cut short without it. Rem. 1 and Prop. 2 follow from Eq. 1 directly, and Prop. 1 is an identity over any world. The remainder are proved in the stylized model $\mathcal{M}_0$ of App. B, which carries the acute extension. With $V_{\mathrm{ind}}(\rho;\theta)$ the paired gain of a protective response ρ on body θ (§3), call θ **at risk** if $\rho_\emptyset$ loses the work before the horizon, and say ρ satisfies **off-risk coupling within** ε if on bodies not at risk it differs from $\rho_\emptyset$ by at most ε.

Remark 1 (Concentration over bodies)**.** *If a protective ρ satisfies off-risk coupling within ε, then $\mathbb{E}_\theta[V_{\mathrm{ind}}] \le q\,\Delta_b + \varepsilon$ over the population, while the conditional value $\mathbb{E}_\theta[V_{\mathrm{ind}} \mid \text{at risk}]$ is the at-risk gain itself, bounded by Δ_b and not diluted by q.*

Corollary 1 (What a mean over bodies answers)**.** *Under the same premise the population mean understates the value to a body that needs it by $1/q$.*

We therefore report the distribution and the conditional value throughout.

Proposition 1 (Schedule confound)**.** *Let σ^* be the best schedule, the best policy reading nothing, and $\chi := \mathbb{E}_\theta[J(\sigma^*) - J(\rho_\emptyset)] \,/\, \mathbb{E}_\theta[V_{\mathrm{ind}}]$ the share of the per-body value it captures. Then $V(C \mid \emptyset) = (1-\chi)\,\mathbb{E}_\theta[V_{\mathrm{ind}}]$, and $V(C \mid \emptyset) \le \mathbb{E}_\theta[V_{\mathrm{ind}}]$ whenever the ablated arm is at least as good in the mean as $\rho_\emptyset$ (proof in App. B).*

The identity holds for any component measured by ablation against a population-trained agent, so an ablation mean is silent wherever a prior can carry the schedule, and we report both quantities in every world.

Proposition 2 (Identification requires wear)**.** *Under Eq. 1 two wear resistances separate only in proportion to the wear already suffered, so every statistic informative about κ is a function of that wear. A body learns its own wear resistance only by spending the integrity the knowledge would have protected (App. B).*

On the channel we ship, no residual information about κ survives in the observation (App. E). Since the proposition is a property of the filtration, it binds every $\mathcal{F}^C$-measurable policy, learned or static, so the trait arrives priced in damage and the value of the channel arises from changed behaviour.

Lemma 1 (Where the constrained agent goes, and why retention is a precondition)**.** *Consider the program of Eq. 3, $\max_a \sum_j a_j \hat{y}_j$ subject to $\sum_j a_j \hat{c}_j \le \tau$, where $\hat{c}_j$ equals the felt cost on the activities the agent has worked and retained and equals the endowment $\hat{c}^0_j = 0$ on all others.* ***(a)*** Escape. *When the constraint binds, the optimum moves the day onto the best-*paid *activity whose cost the agent does not know, whatever that activity truly costs.* ***(b)*** Vacuity without retention. *The cost of an activity is written only after a period that worked it, and each period's allocation is chosen before that period's signal arrives. With zero retention every decision is therefore taken with $\hat{c} \equiv 0$, so the constraint never binds and the agent keeps none of the protective value, not even the $1/R$ share a memoryless agent keeps in the acute case (Thm. 1(b)).*

The proof is given in App. B, where part (b) is Cor. 3. Retention resolves (a) by converting untried into tried one option at a time, and Prop. 12 bounds what a trace can add given what the slow loop installs, and its value vanishes when the endowment is exact.

Proposition 3 (The paying regime, in $\mathcal{M}_0$)**.** *In $\mathcal{M}_0$ under time-additive J, with s the signal, h the memory and w the valuation, if any of the following fails the optimal policy in $\Pi(\{s,h,w\})$ is unprotective and the channels carry no decision-relevant information (Howard, 1966):* ***(i)*** *the threat is individually unpredictable* ex ante*,* ***(ii)*** *it is cheap to avoid* ex post*, the substitution premium c not being bounded away from zero, and* ***(iii)*** *it is expensive to ignore, the value of survived time not being exhausted by a finite milestone schedule. Each is proved separately in Prop. 7 in App. B.*

All three conditions are properties of a world that can be checked before it is run, and we check them in every world we report.

Theorem 1 (Conjunction, acute case)**.** *In $\mathcal{M}_0$ under a law-invariant objective Φ of the per-life return, of which J is one instance, assume (i) and (iii) of Prop. 3, a calibrated gain (Prop. 8), an opaque event, in that no later consequence reveals it, and that mean return alone does not pay protection, $c \geq (p^+ - p)\, L_e$, with p and p^+ the per-step event probability before and after the first event, L_e the event cost and c the substitution premium. Then the Φ-optimal policy in $\Pi(\mathcal{F}^C)$ is protective if and only if (for (b), asymptotically in R)* ***(a)*** sensing*: the event enters the filtration at its time,* ***(b)*** persistence*: it stays there, without which at most a $1/R$ fraction of the gain is captured, R being the steps remaining, and* ***(c)*** tail-sensitivity*: Φ charges an event more than mean return does, by enough that $c < c^*_\Phi := (p^+ - p)\, L^\Phi_e$.*

The theorem is stated for the acute extension of Eq. 1 and proved as Thm. 2 in App. B. Each condition names a role that several modules can fill, (b) a nociceptive memory, recurrence or the action history, and (c) $w\,s$, death or CVaR. Under Eq. 1 the count drops to two, wear supplying (c) through the dynamics (Prop. 4). It is tested in the fleet of §5.1, the one world whose premises we verified. We removed the realization instead, and the fleet then separates on w alone (Thm. 2 states which realizations L^Φ_e admits).

Proposition 4 (Conjunction, chronic case)**.** *In the chronic model $\mathcal{M}_1$ of App. B, a wear budget spent across profiles whose rates are unknown and whose yield falls with the wear,* ***(a)*** *the value of sensing is $\mathrm{EVPI} - \Pi_{\mathrm{probe}}$, the between-body information about which profile this body tolerates, less the output forgone in working each probed profile once.* ***(b)*** *A policy that does not retain its readings pays Π_{probe} and cannot collect the* EVPI*, since the state carries the wear already spent but no comparison* across *profiles.* ***(c)*** *is no longer a condition on Φ.*

There a profile's wear is felt only while it is worked and no event is observed without cost, which is why (a) is net, and the price of (c) lies in the dynamics, so removing it removes the body. The proposition was checked in a separate probe family (App. E).

5 Experiments

Setup. Four worlds carry one stack (Table 1). The *floor layer* is the chronic case, with cartilage wear takes away the deep flexion that kneeling needs, eight day types, a total-output floor at half the healthy output, ages 16 to 65, with 2,000 paired bodies and individual probes re-solving Eq. 3 once a year (App. E). The *care robot* lifts patients under a certification that ends its service (500 paired lives, endpoints certified months and output). The *rover* is the negative control, its wear deterministic in load (500 paired lives). The *latent-defect fleet* is the acute case and the only world whose premises (i)–(iii) we verified (§5.1), with endpoints death rate and J for its trained arms. In every world but the fleet the same three responses run on the same bodies, with two caution arms as contrasts. Trained arms (PPO, App. C) run in every world and carry the full stack $\{s, h, w\}$ and three ablations, clinical cases rather than knockouts: no signal (Charcot's deafferented joint, Nagasako et al., 2003), no weight on pain (asymbolia, Berthier et al., 1988), no retention, with the same observation and the same J, so only the filtration and the forgetting rate differ across arms. Signs and gates were pre-registered (72 declarations).

The population-trained arms are the slow loop, and their contrast measures what a species prior gains (Def. 1). At the anchored cell the channel is worth $+4.4$ years and $+2.4$ output to a body and $+0.65$ years and -0.54 output to the population (24 paired seeds, 20 positive on years, Wilcoxon signed rank $p = 4.4 \times 10^{-5}$, Table 1).

Against the body that works at full effort the apparatus yields $+4.4$ years, 95% CI $[+4.23, +4.53]$, and $+2.4$ output, $[+2.33, +2.53]$, paired over 2,000 bodies (Wilcoxon signed rank, $p < 10^{-15}$ on both), and no row of Table 1 dominates the retaining body. It carries the highest output, the caution rows and the full-stack trained arm obtain their extra years at the cost of output, and the blind trained arm lies below it on both, so the headline is a Pareto statement. We report a corner and a band, where the corner is the grounded sensitization gain, the rate at which the felt signal rises with accumulated wear ($\alpha_{\mathrm{sens}} = 8.51$, App. E), at the measured kneeling share (41.0% of the day, SD 7.5, Jensen et al., 2010), the least favourable corner on years we measured. Across the anchored span of that gain (2.45 to 8.51) the gain runs $+4.4$ to $+5.8$ years and $+1.9$ to $+2.4$ output, with no body losing at either end, and over the set spread of wear resistance it runs $+3.1$ to $+5.6$ years (App. E). The present-signal row and the retaining row differ only in what they retain, so their difference isolates

	floor layer, absolute				care robot		rover		fleet	
what the body does	works to	output	D_{phys}	M_{phys}	life	output	life	output	life	output
works at full effort, reads nothing	55.2	33.7	0.227	0.739	51.5	36.9	120.1	96.4	74.9	67.6
spares from the first day[†]	61.5	34.4	**0.188**	**0.780**	+226%	+121%	+151%	+101%	+15%	−10%
fixed endowment $\alpha{=}2$[†§]	**61.6**	35.3	0.191	0.775	—	—	—	—	—	—
trained, feels nothing[‡]	59.2	33.4	0.200	0.768	+67%	+30%	−30%	−46%	−6%	−26%
trained, full stack[‡]	59.9	32.9	0.191	0.778	+679%	+432%	+26%	+2%	+26%	+20%
reads the present signal, no retention	56.2	34.3	0.223	0.743	+528%	+309%	+163%	+108%	+1%	+1%
reads the signal and remembers	59.6	**36.1**	0.205	0.758	+462%	+269%	+165%	+110%	$\geq$ +34%	+34%

Table 1: **Three ways to use one body, in four worlds, on the same individuals.** All arms are paired, and no arm carries a schedule learned across lives (Def. 1). Floor layer: *works to* is the age the trade ends, *output* career output, D_{phys} cartilage lost, M_{phys} meniscal competence retained (2,000 bodies). Machines: the first row gives the service life (certified months for the care robot, months otherwise) and the output of the arm that reads nothing (500 paired lives per robot, 5,000 in the fleet), the other rows the percent change against it. The fleet runs to a horizon of 100 months, which the retaining arm reaches in every life, so its gain is a lower bound. [†]Not admissible as a reference, since it already knows which work is gentlest. [§]Trade only, since a machine reads every activity at once and an endowment is overwritten by the first reading. [‡]PPO. Trade: 10^6 environment steps, 24 paired seeds, 50,000 paired lives per seed, output as the trade paid it. Robots: 3×10^6 steps, 8 seeds, the rover at residual stream 0.10. Fleet: 10^6 steps, 18 seeds, entropy 0 (App. C). The blind trained arm already carries the population's schedule (it reaches 59.2 without a channel), so these rows measure what a species prior gains from the channel, while the value to one body is the last row against the first.

condition (b) of Thm. 1, persistence, worth three to four years in the trade and negative on the care robot (§5.3).

The reference carries no schedule. The claim is what one life adds to a body that knows nothing, and the reference is chosen for that question. Sparing from the first day works to 61.5 years and the best schedule fitted across bodies to 61.6, against the retaining body's 59.6, and a reader may take either for the fairer reference. Both encode which work is gentlest, knowledge an individual entering the trade does not hold, as Charcot's patient does not, who carries the species prior and still reloads the joint. What a schedule captures is therefore reported beside the claim, $\chi = 1.40$ on years (Prop. 1), together with the channel's value beyond the best schedule, -1.77 predicted and $+0.65$ measured, which is the second question of §3.

5.1 A LEARNER ACQUIRES THE CONJUNCTION WHERE THE REGIME HOLDS

The regime map, the three conditions of Prop. 3, made one out-of-sample prediction, filed before the world was built, that where (i)–(iii) hold the conjunction must return. A fleet of machines in the field chooses among eight near-equivalent classes of work (ii). Which class a given machine cannot tolerate is drawn per unit and is unidentifiable before the first incident (i). Incidents compound and the third writes the machine off (iii). The task economy is constructed, the event statistics come from field data on 10^5-drive populations (Pinheiro et al., 2007; Schroeder & Gibson, 2007). The prediction held, in that at the deterministic optimum (entropy 0, 18 seeds per arm) the full stack writes off 0.15 of its machines at $J = 0.79$, without the signal 0.55 at $J = 0.49$, without retention 0.53 at $J = 0.50$ (seed-level Mann–Whitney on the death rate, Cliff's $\delta = -0.85$ and -0.83, $p = 1.4 \times 10^{-5}$ and 2.1×10^{-5}, and on J the seed distributions do not overlap, App. C.1). The damage cost is inert, 0.09 at $J = 0.79$ ($\delta = +0.03$, $p = 0.89$), because the world already prices damage through the loss of the work (App. C.1). One pre-registered element failed and is reported. The absolute protection level (median death rate below 0.15) is missed at the default entropy (median 0.38) and reached by 8 of 18 seeds at entropy 0 (median 0.154), while the contrasts hold at both settings (App. C.1). A probe that reads the signal and retains nothing past the day gains 1% of service life over the blind probe, the one-step reallocation that condition (b) predicts, against the 34% of the retaining probe (Table 1).

5.2 A LONGER WORKING LIFE AND MORE OUTPUT, AND THE MEMORY CARRIES THEM

On one floor layer's knee, feeling the wear lengthens the working life and raises its output, and retention carries almost all of it (Table 1, Fig. 3 in App. E). Over $2{,}000$ paired bodies 69.3% gain years and **none lose** any, the median gain being 5 years. On output 76.9% gain and 23.1% lose. The premise of Rem. 1 holds here in one coordinate only. A body not at risk works to the horizon in both arms, so on career length $\varepsilon = 0$ is forced by the cap, while in output, which nothing caps, those bodies lose 0.40 at the observed mix and gain 0.18 where specialization is free. With $q = 0.693$ the bound is exact in years and within 0.4 in output (Cor. 2 in App. B), so the mean of $+4.4$ years is q times the $+6.3$ an at-risk body experiences. The split is stable in its order and varies in its level. The memoryless arm lies on a degenerate vertex of its program and spans 55.7 to 56.2 over random tie-breaks, so retention carries $+3.4$ to $+3.9$ of the $+4.4$ years and sensing alone $+0.5$ to $+1.0$ (App. E). Specialization moves the gain from $+4.38$ to $+6.03$ years, the permissive end ending at 61.2 with 37.9 in output. A body that retains nothing at all, not even the last period's reading, lays floors in every working year, its felt constraint never binding, and the signal is then worth $+0.00$ years (Cor. 3).

What the trace provides is an end to resuming work the body has already left. Reading only the present signal it resumes the heavy work 6.6 times in one working life, and with persistence never (Table 5). A falling kneeling share does not by itself indicate substitution. The body that reads nothing is removed from the work by the loss of capability and lays carpet, its day ending on two activities against four (Fig. 4 in App. E). Feeling alone reaches the gentle work in 1% of its last working days, the retaining body in 36%, so its substitution is better aimed.

5.3 ONE ROBOTIC WORLD PAYS AND THE OTHER DOES NOT

Both machines share the stack above the afferent map α and wear by Palmgren–Miner (Miner, 1945; Lundberg & Palmgren, 1947) on MuJoCo-measured load spectra. The care robot sextuples its certified service life and quadruples its output ($51.5 \to 323.3$ months, $36.9 \to 150.8$, Table 1), paired over 500 lives with every life gaining (intervals in App. C). Here *retention does not pay*. A trace returns 34.2 of those months, $[29.3,\ 39.1]$, and 14.7 of that output, $[12.5,\ 16.9]$, because the robot reads its cost off motor-level signals through a known load matrix, so the premise of Lemma 1 fails and a trace has nothing to aim (Prop. 12). In the rover the threat is deterministic in load, so (i) fails and blind caution acts on the same lever. Caution alone yields 181.2 of the 195.8 months, and the signal adds $+14.6$ months and $+7.3$ output (Prop. 3(i)).

Three limits that change how a result is read. Of the ten limits in App. A, three bear on the numbers above: the modelled day is more knee-loaded than the measured one, capping that share reverses the sign (the cap is an institution and marks the map's boundary), and only $\Delta\theta = 18^\circ$ passes the register gate (App. A, last paragraph).

6 CONCLUSIONS AND FUTURE WORK

A feeling body stays intact by a conjunction of sensing, retention and an intrinsic damage cost, which pays where threats are unpredictable, avoidable and unforgiving. It provides **resilience**, the work held, and **body-sparing**, the load shed, worth on one body $+4.4$ years and $+2.4$ output at the observed task mix, almost all of it from retention. On working life two of three bodies gain and none lose, on output 77% gain and 23% lose. Three things follow beyond this paper. Agents deployed once in a body they cannot learn about are the general case for robots, prostheses and workers, and epoch one is the setting in which to design them. The conditions of Prop. 3 can be checked before a world is built, and the fleet shows that a learner acquires the conjunction where they hold. Prop. 1 applies to every component valued by ablation against a population-trained agent, and one paired run supplies what the ablation omits. Open are the slow loop itself, which phylogeny a memory needs and whether a recurrent policy can learn the latch it here carries, and the worlds beyond one trade and one joint. The ways a protection experiment fails, and the checks that catch them, are in App. D.

REPRODUCIBILITY STATEMENT

Every number printed in this paper is recomputed from the file it came from by a checker that ships with the code, `experiments/paper_zahlen.py`: it reads the result files, recomputes each quantity, compares it against the printed value, and lists separately any printed quantity for which it has no source, so that an unsourced number is visible rather than silent. A test fails when the two drift. The supplementary archive contains the drivers and environments, the 72 pre-registration files under `experiments/decl/` that fix the sign and the gate of each run before it is submitted, the null-gate tests, and the result files behind every table and figure (695 files, 14 MB compressed, built by `experiments/supplement_bauen.py`, which runs the checker and the null-gate tests inside the unpacked archive). What it does not contain is the per-life record of every run: those are 21 GB, and each is regenerated by the driver named in the caption of the table that uses it. Where a run could not be located in this archive we removed the number rather than attribute it: two columns of Table 2, one fleet table and two blocks of trained figures were withdrawn on that ground, and every quantity that remains is bound to a file. The four worlds, their fixed probes as baselines and the checker form one suite in which any cell of Table 1 is one command (`experiments/bodenleger_lp_cell.py` for the trade, `pflege_familie.py`, `rover_familie.py` and `hh_familie.py` for the machines), so that a new response or a new world is scored against the same probes. The care-robot column was later restored from a new run, and the rover column is omitted because it carries no information (App. C). A calibration loss quoted in App. B without a locatable run was removed on the same ground and the statement kept qualitative.

Where each table comes from. Running text is pinned number by number. Tables are bound to the file that reproduces them, and the unnumbered tabulars of App. B are introduced by the text that names their driver. Table 1: `results/bodenleger_lp_b042`, `results/pareto_vier*.json`, `results/pflege_familie`. The machine intervals of §5.3: `results/pflege_familie/familie_ci.json` and `results/rover_familie`. Table 2: `results/tab_leiter_spalten.json`, the per-seed aggregate of `results/knee_fdecke2_hpc` (floor-layer column, whose per-seed summaries are not in the archive, only their aggregate), of `results/hh_verankert2_hpc` (latent-defect column) and of `results/pflege_boden_hpc` (care-robot column), per-seed summaries included. The trained machine rows of Table 1: `results/tab1_maschinen_trainiert.json` from `results/pflege_boden_hpc`, `results/rover_boden_hpc` and `results/hh_verankert2_hpc`. Table 3: `results/hh_verankert2_hpc`. The fleet's probe family: `results/hh_familie`. Table 5: `results/substitution_spur_b042.json`. Fig. 2: `results/fig_arme_welten.json` (`experiments/fig_arme_welten.py`, from the files of Table 1, `results/rover_familie` and `results/hh_familie`). The trained rows of Table 1: `results/knee_mischung_hpc`, their D_{phys} and M_{phys} from `results/knee_mischung_phys.json` (the same 50,000 paired lives per seed, re-evaluated from the checkpoints, `results/knee_mischung_phys_hpc`). Table 6: `results/bodenleger_lp_b042`. Table 7: `results/knee_individuum_sweep_v5`, `results/hh_familie`, `results/pflege_familie`, `results/rover_familie`.

ETHICS STATEMENT

This work uses no human subjects, no personal data and no data collected by the authors. The biological anchors are published summary statistics from occupational-medicine and biomechanics literature, cited at each use. The object of study is a simulated agent, and the intended reading of its results is about *machines that wear out* and about the design of protection experiments. The floor layer's world is a modelling device anchored on published exposure and epidemiology, not a claim about any individual worker. We note one risk of the framing explicitly: a result stating when a body model earns its keep can be read as a statement about which workers are worth protecting. It is not one. The model measures what an agent must carry to remain competent in one life, and its clearest boundary condition points the other way, in App. A(10): where a work rule caps exposure, the protection the body model buys is not needed.

THE USE OF LARGE LANGUAGE MODELS

A large language model was used as a coding and writing assistant throughout: it wrote and refactored parts of the experiment code, ran and inspected the analyses under the authors' direction, drafted and edited prose, and audited the manuscript against the result files, which is how several of the defects reported in App. D were found. It is not an author, it contributed no research idea or claim of its own, and every number, anchor and conclusion in this paper was checked against its source by the authors. The checker described in the Reproducibility Statement exists because assistant-drafted text can carry numbers that no run supports.

REFERENCES

Joshua Achiam, David Held, Aviv Tamar, and Pieter Abbeel. Constrained policy optimization. In *International Conference on Machine Learning*, 2017.

Mohammed Alshiekh, Roderick Bloem, Rüdiger Ehlers, Bettina Könighofer, Scott Niekum, and Ufuk Topcu. Safe reinforcement learning via shielding. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2018.

Eitan Altman. *Constrained Markov Decision Processes*. Chapman and Hall/CRC, 1999.

Aaron D. Ames, Samuel Coogan, Magnus Egerstedt, Gennaro Notomista, Koushil Sreenath, and Paulo Tabuada. Control barrier functions: Theory and applications. In *18th European Control Conference (ECC)*, pp. 3420–3431, 2019.

Charalampos P. Andriotis and Konstantinos G. Papakonstantinou. Managing engineering systems with large state and action spaces through deep reinforcement learning. *Reliability Engineering & System Safety*, 191:106483, 2019.

Karl Johan Åström. Optimal control of Markov processes with incomplete state information. *Journal of Mathematical Analysis and Applications*, 10(1):174–205, 1965.

Peter Auer, Nicolò Cesa-Bianchi, and Paul Fischer. Finite-time analysis of the multiarmed bandit problem. *Machine Learning*, 47(2–3):235–256, 2002.

Gábor Bartók, Dean P. Foster, Dávid Pál, Alexander Rakhlin, and Csaba Szepesvári. Partial monitoring—classification, regret bounds, and algorithms. *Mathematics of Operations Research*, 39(4):967–997, 2014.

Jakob Bauer, Kate Baumli, Feryal Behbahani, Avishkar Bhoopchand, Nathalie Bradley-Schmieg, Michael Chang, Natalie Clay, Adrian Collister, Vibhavari Dasagi, Lucy Gonzalez, et al. Human-timescale adaptation in an open-ended task space. In *International Conference on Machine Learning*, 2023.

Nicole Bäuerle and Jonathan Ott. Markov decision processes with average-value-at-risk criteria. *Mathematical Methods of Operations Research*, 74(3):361–379, 2011.

Pablo Bellosta-López, Víctor Domentín-Casanova, et al. Long-term consistency of clinical sensory testing measures for pain assessment. *The Korean Journal of Pain*, 36(2):173–183, 2023. doi: 10.3344/kjp.23011.

M.-J. Berthiaume, J.-P. Raynauld, J. Martel-Pelletier, F. Labonté, G. Beaudoin, D. A. Bloch, D. Choquette, B. Haraoui, R. D. Altman, M. Hochberg, J. M. Meyer, G. A. Cline, and J.-P. Pelletier. Meniscal tear and extrusion are strongly associated with progression of symptomatic knee osteoarthritis as assessed by quantitative magnetic resonance imaging. *Annals of the Rheumatic Diseases*, 64(4):556–563, 2005. PMID 15374855.

Marcelo Berthier, Sergio Starkstein, and Ramon Leiguarda. Asymbolia for pain: a sensory-limbic disconnection syndrome. *Annals of Neurology*, 24(1):41–49, 1988.

John R. Birge and François Louveaux. *Introduction to Stochastic Programming*. Springer, 2nd edition, 2011.

Charles Blundell, Benigno Uria, Alexander Pritzel, Yazhe Li, Avraham Ruderman, Joel Z. Leibo, Jack Rae, Daan Wierstra, and Demis Hassabis. Model-free episodic control. *arXiv preprint arXiv:1606.04460*, 2016.

Károly Boda and Jerzy A. Filar. Time consistent dynamic risk measures. *Mathematical Methods of Operations Research*, 63(1):169–186, 2006.

Mark E. Bouton. Context and behavioral processes in extinction. *Learning & Memory*, 11, 2004.

Ronen I. Brafman and Moshe Tennenholtz. R-max: a general polynomial time algorithm for near-optimal reinforcement learning. *Journal of Machine Learning Research*, 3:213–231, 2002.

Archana Bura, Aria HasanzadeZonuzy, Dileep Kalathil, Srinivas Shakkottai, and Jean-François Chamberland. DOPE: Doubly optimistic and pessimistic exploration for safe reinforcement learning. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2022.

Annie S. Chen, Archit Sharma, Sergey Levine, and Chelsea Finn. You only live once: Single-life reinforcement learning. In *Advances in Neural Information Processing Systems*, 2022.

Yinlam Chow, Aviv Tamar, Shie Mannor, and Marco Pavone. Risk-sensitive and robust decision-making: a CVaR optimization approach. *Advances in Neural Information Processing Systems*, 2015.

A. D. Craig. How do you feel? interoception: the sense of the physiological condition of the body. *Nature Reviews Neuroscience*, 3, 2002.

Antoine Cully, Jeff Clune, Danesh Tarapore, and Jean-Baptiste Mouret. Robots that can adapt like animals. *Nature*, 521(7553):503–507, 2015.

D. M. Ditchen, R. P. Ellegast, T. Gawliczek, B. Hartmann, and M. A. Rieger. Occupational kneeling and squatting: development and validation of an assessment method combining measurements and diaries. *International Archives of Occupational and Environmental Health*, 88(2):153–165, 2015a. PMID 24859645.

Dirk M. Ditchen, Rolf P. Ellegast, Bernd Hartmann, and Monika A. Rieger. Occupational kneeling and squatting: development and validation of an assessment method combining measurements and diaries. *International Archives of Occupational and Environmental Health*, 88(2):153–165, 2015b.

Yan Duan, John Schulman, Xi Chen, Peter L. Bartlett, Ilya Sutskever, and Pieter Abbeel. RL^2: Fast reinforcement learning via slow reinforcement learning. *arXiv preprint arXiv:1611.02779*, 2016.

Michael O. Duff. *Optimal Learning: Computational Procedures for Bayes-Adaptive Markov Decision Processes*. PhD thesis, University of Massachusetts Amherst, 2002.

Yonathan Efroni, Shie Mannor, and Matteo Pirotta. Exploration-exploitation in constrained MDPs. *arXiv preprint arXiv:2003.02189*, 2020.

Martin Englund, Ewa M. Roos, and L. Stefan Lohmander. Impact of type of meniscal tear on radiographic and symptomatic knee osteoarthritis: a sixteen-year followup of meniscectomy with matched controls. *Arthritis & Rheumatism*, 48(8):2178–2187, 2003. doi: 10.1002/art.11088.

Tohru Fukubayashi and Hisashi Kurosawa. The contact area and pressure distribution pattern of the knee. *Acta Orthopaedica Scandinavica*, 51, 1980.

Javier García and Fernando Fernández. A comprehensive survey on safe reinforcement learning. *Journal of Machine Learning Research*, 16, 2015.

Mohammad Ghavamzadeh, Shie Mannor, Joelle Pineau, and Aviv Tamar. Bayesian reinforcement learning: A survey. *Foundations and Trends in Machine Learning*, 8(5-6):359–483, 2015.

Nathan Grinsztajn, Johan Ferret, Olivier Pietquin, Philippe Preux, and Matthieu Geist. There is no turning back: A self-supervised approach for reversibility-aware reinforcement learning. In *Advances in Neural Information Processing Systems*, 2021.

Monika I. Hasenbring and Jeanine A. Verbunt. Fear-avoidance and endurance-related responses to pain: new models of behavior and their consequences for clinical practice. *The Clinical Journal of Pain*, 26(9):747–753, 2010. PMID 20664333.

M. A. Hassaballa, A. J. Porteous, J. H. Newman, and C. A. Rogers. Can knees kneel? Kneeling ability after total, unicompartmental and patellofemoral knee arthroplasty. *The Knee*, 10(2):155–160, 2003. PMID 12787999.

Ronald A. Howard. Information value theory. *IEEE Transactions on Systems Science and Cybernetics*, 2, 1966.

Louis S. Jacobson, Robert J. LaLonde, and Daniel G. Sullivan. Earnings losses of displaced workers. *American Economic Review*, 83, 1993.

Andrew K. S. Jardine, Daming Lin, and Dragan Banjevic. A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing*, 20(7):1483–1510, 2006.

Bengt Järvholm, Mikael Stattin, Suzan J. W. Robroek, Urban Janlert, Björn Karlsson, and Alex Burdorf. Heavy work and disability pension: a long term follow-up of Swedish construction workers. *Scandinavian Journal of Work, Environment & Health*, 40(4):335–342, 2014. doi: 10.5271/sjweh.3413.

Lilli Kirkeskov Jensen. Knee osteoarthritis: influence of work involving heavy lifting, kneeling, climbing stairs or ladders, or kneeling/squatting combined with heavy lifting. *Occupational and Environmental Medicine*, 65(2):72–89, 2008. doi: 10.1136/oem.2007.032466.

Lilli Kirkeskov Jensen, Soren Rytter, and Jens Peter Bonde. Exposure assessment of kneeling work activities among floor layers. *Applied Ergonomics*, 41(2):319–325, 2010. PMID 19766986.

Abbas Kazerouni, Mohammad Ghavamzadeh, Yasin Abbasi-Yadkori, and Benjamin Van Roy. Conservative contextual linear bandits. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2017.

Mehdi Keramati and Boris Gutkin. Homeostatic reinforcement learning for integrating reward collection and physiological stability. *eLife*, 3, 2014.

Khimya Khetarpal, Matthew Riemer, Irina Rish, and Doina Precup. Towards continual reinforcement learning: A review and perspectives. *Journal of Artificial Intelligence Research*, 75:1401–1476, 2022.

Georgia-Ann Klutke, Peter C. Kiessler, and Martin A. Wortman. A critical look at the bathtub curve. *IEEE Transactions on Reliability*, 52(1):125–129, 2003.

Johannes Kuehn and Sami Haddadin. An artificial robot nervous system to teach robots how to feel pain and reflexively react to potentially damaging contacts. *IEEE Robotics and Automation Letters*, 2(1):72–79, 2017.

Ashish Kumar, Zipeng Fu, Deepak Pathak, and Jitendra Malik. RMA: Rapid motor adaptation for legged robots. In *Robotics: Science and Systems*, 2021.

Michael Laskin, Luyu Wang, Junhyuk Oh, Emilio Parisotto, Stephen Spencer, Richie Steigerwald, DJ Strouse, Steven Hansen, Angelos Filos, Ethan Brooks, Maxime Gazeau, Himanshu Sahni, Satinder Singh, and Volodymyr Mnih. In-context reinforcement learning with algorithm distillation. In *International Conference on Learning Representations*, 2023.

Jonathan N. Lee, Annie Xie, Aldo Pacchiano, Yash Chandak, Chelsea Finn, Ofir Nachum, and Emma Brunskill. Supervised pretraining can learn in-context reinforcement learning. In *Advances in Neural Information Processing Systems*, 2023.

Zachary C. Lipton, Kamyar Azizzadenesheli, Abhishek Kumar, Lihong Li, Jianfeng Gao, and Li Deng. Combating reinforcement learning's Sisyphean curse with intrinsic fear. *arXiv preprint arXiv:1611.01211*, 2016.

Tao Liu, Ruida Zhou, Dileep Kalathil, P. R. Kumar, and Chao Tian. Learning policies with zero or bounded constraint violation for constrained MDPs. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2021.

Gustaf Lundberg and Arvid Palmgren. Dynamic capacity of rolling bearings. *Acta Polytechnica, Mechanical Engineering Series*, 1, 1947.

Albert Madansky. Inequalities for stochastic linear programming problems. *Management Science*, 6(2):197–204, 1960.

Catherine Mailloux, Louis-David Beaulieu, Timothy H. Wideman, and Hugo Massé-Alarie. Within-session test-retest reliability of pressure pain threshold and mechanical temporal summation in healthy subjects. *PLoS ONE*, 16(1):e0245278, 2021. doi: 10.1371/journal.pone.0245278.

Kingson Man and Antonio Damasio. Homeostasis and soft robotics in the design of feeling machines. *Nature Machine Intelligence*, 1:446–452, 2019.

Anna Marcuzzi, Paul J. Wrigley, et al. The long-term reliability of static and dynamic quantitative sensory testing in healthy individuals. *Pain*, 158(7):1217–1223, 2017. doi: 10.1097/j.pain.0000000000000901.

Susanne Maschek, Wolfgang Wirth, Christoph Ladel, Marie-Pierre Hellio Le Graverand, and Felix Eckstein. Rates and sensitivity of knee cartilage thickness loss in specific central reading radiographic strata from the osteoarthritis initiative. *Osteoarthritis and Cartilage*, 22(10):1550–1553, 2014. doi: 10.1016/j.joca.2014.05.015. PMID 25278063.

William Q. Meeker and Luis A. Escobar. *Statistical Methods for Reliability Data*. Wiley, 1998.

Milton A. Miner. Cumulative damage in fatigue. *Journal of Applied Mechanics*, 12, 1945.

Elna M. Nagasako, Anne Louise Oaklander, and Robert H. Dworkin. Congenital insensitivity to pain: an update. *Pain*, 101(3):213–219, 2003.

M. Opara Zupancic and N. Sarabon. Disease severity affects knee range of motion but not strength deficits in knee osteoarthritis: a systematic review and meta-analysis. *Frontiers in Medicine*, 13, 2026. doi: 10.3389/fmed.2026.1737973.

Eduardo Pinheiro, Wolf-Dietrich Weber, and Luiz André Barroso. Failure trends in a large disk drive population. In *5th USENIX Conference on File and Storage Technologies (FAST)*, pp. 17–28, 2007.

Alexander Pritzel, Benigno Uria, Sriram Srinivasan, Adrià Puigdomènech Badia, Oriol Vinyals, Demis Hassabis, Daan Wierstra, and Charles Blundell. Neural episodic control. In *International Conference on Machine Learning*, 2017.

Kate Rakelly, Aurick Zhou, Chelsea Finn, Sergey Levine, and Deirdre Quillen. Efficient off-policy meta-reinforcement learning via probabilistic context variables. In *International Conference on Machine Learning*, 2019.

Lee C. Rogers, Robert G. Frykberg, David G. Armstrong, Andrew J. M. Boulton, Michael Edmonds, George Ha Van, Agnès Hartemann, Frances Game, William Jeffcoate, Alexandra Jirkovska, Edward Jude, Stephan Morbach, William B. Morrison, Michael Pinzur, Dario Pitocco, Lee Sanders, Dane K. Wukich, and Luigi Uccioli. The Charcot foot in diabetes. *Diabetes Care*, 34(9):2123–2129, 2011.

H. Roos, M. Laurén, T. Adalberth, E. M. Roos, K. Jonsson, and L. S. Lohmander. Knee osteoarthritis after meniscectomy: prevalence of radiographic changes after twenty-one years, compared with matched controls. *Arthritis & Rheumatism*, 41(4):687–693, 1998. PMID 9550478.

Michael Roth, Wolfgang Wirth, Katja Emmanuel, Adam G. Culvenor, and Felix Eckstein. The contribution of 3d quantitative meniscal and cartilage measures to variation in normal radiographic joint space width. *European Journal of Radiology*, 87:90–98, 2017. PMID 28065381.

Abhinav Saxena, Kai Goebel, Don Simon, and Neil Eklund. Damage propagation modeling for aircraft engine run-to-failure simulation. In *International Conference on Prognostics and Health Management (PHM)*, 2008.

Bianca Schroeder and Garth A. Gibson. Disk failures in the real world: What does an MTTF of 1,000,000 hours mean to you? In *5th USENIX Conference on File and Storage Technologies (FAST)*, pp. 1–16, 2007.

John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.

Xiao-Sheng Si, Wenbin Wang, Chang-Hua Hu, and Dong-Hua Zhou. Remaining useful life estimation: A review on the statistical data driven approaches. *European Journal of Operational Research*, 213(1):1–14, 2011.

Jayakumar Subramanian, Amit Sinha, Raihan Seraj, and Aditya Mahajan. Approximate information state for approximate planning and reinforcement learning in partially observed systems. *Journal of Machine Learning Research*, 23(12):1–83, 2022.

Hidenori Suzuki, Shu Tahara, Mao Mitsuda, Masahiro Funaba, Kazuhiro Fujimoto, Hiroaki Ikeda, Hironori Izumi, Kiminori Yukata, Kazushige Seki, Kenji Uranami, Kiyoshi Ichihara, Norihiro Nishida, and Takashi Sakai. Reference intervals and sources of variation of pressure pain threshold for quantitative sensory testing in a Japanese population. *Scientific Reports*, 13:13043, 2023. doi: 10.1038/s41598-023-40201-w.

Aviv Tamar, Yonatan Glassner, and Shie Mannor. Optimizing the CVaR via sampling. In *AAAI Conference on Artificial Intelligence*, 2015.

Philip S. Thomas, Georgios Theocharous, and Mohammad Ghavamzadeh. High confidence off-policy evaluation. In *AAAI Conference on Artificial Intelligence*, 2015.

Josh Tobin, Rachel Fong, Alex Ray, Jonas Schneider, Wojciech Zaremba, and Pieter Abbeel. Domain randomization for transferring deep neural networks from simulation to the real world. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2017.

Matteo Turchetta, Felix Berkenkamp, and Andreas Krause. Safe exploration in finite Markov decision processes with Gaussian processes. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2016.

Alexander Matt Turner, Dylan Hadfield-Menell, and Prasad Tadepalli. Conservative agency via attainable utility preservation. In *AAAI/ACM Conference on AI, Ethics, and Society*, 2020.

Kelsey J. Vazquez, Jacob T. Andreae, and Corinne R. Henak. Cartilage-on-cartilage cyclic loading induces mechanical and structural damage. *Journal of the Mechanical Behavior of Biomedical Materials*, 98:262–267, 2019. doi: 10.1016/j.jmbbm.2019.06.023. PMID 31280053. Until 2026-09-16 this entry was keyed and authored as "Bonnevie 2019" with an invented title; the PMID always pointed here. The 4.35–4.73 MPa damage-initiation figure cited from it is model-derived and remains a soft anchor (see prd_kappa_safe_provenance_diagnosis.md).

Akifumi Wachi and Yanan Sui. Safe reinforcement learning in constrained Markov decision processes. In *International Conference on Machine Learning (ICML)*, 2020.

Peter Whittle. Restless bandits: activity allocation in a changing world. *Journal of Applied Probability*, 25(A):287–298, 1988.

Clifford J. Woolf. What is this thing called pain? *Journal of Clinical Investigation*, 120(11):3742–3744, 2010.

Yifan Wu, Roshan Shariff, Tor Lattimore, and Csaba Szepesvári. Conservative bandits. In *International Conference on Machine Learning (ICML)*, 2016.

J. Zelle, M. Barink, R. Loeffen, M. De Waal Malefijt, and N. Verdonschot. Thigh-calf contact force measurements in deep knee flexion. *Clinical Biomechanics*, 22(7):821–826, 2007. PMID 17512647.

Ge Zhou, Minwei Zhao, Xiaoxiao Wang, Xiao Geng, et al. Demographic and radiographic factors for knee symptoms and range of motion in patients with knee osteoarthritis: a cross-sectional study in Beijing, China. *BMC Musculoskeletal Disorders*, 24, 2023. doi: 10.1186/s12891-023-06432-8.

A SCOPE OF THE THEORY: WHAT THE RESULTS DO NOT COVER

(1) Rem. 1 is claimed for the chronic regime, where the premise holds exactly ($\varepsilon = 0.0000$, Cor. 2 in App. B). Under an acute hazard the at-risk set is defined by the event and the coupling is no longer zero, running from -0.0182 to $+0.0181$ over four conditions of the acute-tear knee, and at one of them it is the whole effect. We state the theorem over bodies for that reason and report the acute constants rather than assume them. **(2)** Cor. 1 says what a mean *means*, not that the effect is hard to see. In the trade 28 bodies suffice unpaired. **(3)** Prop. 2 closes the latent trait, not *filtering*: acting on s_t estimates a state. Its chronic half is closed by measurement, and that measurement is a property of the anchored transduction rather than of the regime, since without gain and noise the same observation recovers κ at $R^2 = 0.63$ to 0.75 and $\sigma_s = 0.05$ against our anchored 0.10 already removes it (Prop. 6 in App. B). A finer channel would put this setting back into the estimation regime. **(4)** "Therefore memorization" is an inference, not a theorem: the bound does not close a method adapting on the reward stream (§2). **(5)** Prop. 3 proves the three conditions separately and claims no joint sufficiency. Membership is necessary and not sufficient. **(6)** Thm. 1(a)'s necessity presumes **event opacity**, that no later consequence reveals the event. Where a blind arm reads it off its own falling output, as ours does, (a) bounds only the gain beyond that channel, worth 0.3 years in the trade (Thm. 2 in App. B). **(7)** (b)'s equivalence is asymptotic in R, and an ablation that deletes a memory module tests (b) only where nothing else in the observation carries the event forward. That holds in the fleet. In the knee the observation is afferent-only, so the tissue readings go with the trace, but a five-year rate and a five-year running maximum of the felt signal remain: persistence with a horizon, which is neither a clean removal nor a clean instance. *This is a statement about what an ablation measures*, not about whether a monotone substrate makes retention unnecessary. The pre-registered test of the latter is in §5.3 and went the other way. The configuration that would make the knee a clean test is named in Thm. 2 in App. B and is now wired. The run returned and moved the endpoint against the theory (App. E). **(8)** (c) is close to definitional, since L_e^{Φ} is what Φ charges for an event. The content is the case analysis in Thm. 2 in App. B, which covers the realizations whose effective event cost is fixed within a life and excludes the one, CVaR_β, whose tail set moves with the optimizer. Finally, Thm. 1 is tested in one world: the trade is outside its scope rather than a weaker test of it, since that world fails (iii) for one body, and what carries the headline years there is Lemma 2 with Cor. 3. Pessimism on untried costs does not foreclose that substitution: over the anchored range of $\hat{c}^0$ there is no foreclosure anywhere (App. 12), because the trade always keeps one profile the pessimist still believes he can perform. What the band does show is that the endowment substitutes for retention and not for the signal. **(9)** One limit is not about the theory but about the size of the effect it describes: the wear rate is anchored on people who spare their joints (Maschek et al., 2014; Roth et al., 2017) and applied to a body that never spares, so the anchor already holds part of the answer, in the direction that shrinks what we report. **(10)** The sharpest bound is not a modelling choice but a question about the world: *whether the reference body is allowed to protect itself.* The field anchor enters as a lower bound on laying work, because that is what the trade supplies. Imposing it as an upper bound as well, a work rule capping exposure for every body, reverses the headline. Over $2{,}000$ paired bodies per variant:

what bounds the kneeling share	V_{ind} years	share losing	kneeling share reached
supply only (lower bound, canon)	$\mathbf{+4.38}$	0.0%	0.47
supply only, Ditchen screed share	$+3.18$	0.0%	0.45
supply and a work rule (upper bound too)	$\mathbf{-1.02}$	$\mathbf{43.6}\%$	0.43

It is the first configuration in this paper in which any body loses, and we report it without adopting it, for a reason older than the measurement: the measured 41% is what floor layers *do*, hence the outcome of the mechanism under test, so laying it on the arm that cannot feel hands that arm the fruit of that mechanism, which is the circularity Def. 1 already excludes for a sparing schedule learned across lives. The line we draw is between what the world *offers* (a lower bound, admissible) and

what the worker *chooses* (an upper bound, not admissible as a world constraint). The measurement is therefore not a caveat but the precise statement of where the result lives: *where a work rule caps exposure, a body model is not needed* for what we measure here. Weighting each activity by its knee-loaded time share (Ditchen et al., 2015b; Jensen et al., 2010) gives 61.8% of the day for the body that senses nothing, 49.3% for the one that feels without retaining and 49.6% with the apparatus, against 41.0% (SD 7.5) in the field. Source `results/kniezeit_band.json`, pre-registration `experiments/decl/kniezeit_band.py`.

B STYLIZED MODEL AND PROOFS

Model $\mathcal{M}_0$. A life has H steps. Actions $a \in \{\text{heavy}, \text{light}\}$ yield per-step value y_H and $y_L = y_H - c$ with substitution premium $c \geq 0$. Under heavy, an event occurs w.p. p per step before the first event and $p^+ = \nu p$ ($\nu > 1$, the cascade) after it, while light carries only the baseline hazard p. An event costs L_e (damage, incident cost). In existential variants a second event ends the value stream. All quantities are bounded, and $pH \ll 1$.

Proposition 5 (Concentration over bodies, Rem. 1 of the main text)**.** *Let ρ and $\rho_\emptyset$ be two responses, let B be the set of bodies on which $\rho_\emptyset$ loses the work before the horizon, and $q := \Pr[\theta \in B]$. If the two agree to within ε on $\theta \notin B$, then $\mathbb{E}_\theta[V_{\text{ind}}] \leq q\,\Delta_b + (1-q)\,\varepsilon$ with $\Delta_b := \sup_{\theta \in B} V_{\text{ind}}(\rho; \theta)$, while the conditional value is the at-risk difference by definition.*

Proof. Couple the two responses on the same body and the same draws and split the expectation over $\mathbb{I}\{\theta \in B\}$: off B the integrand is at most ε, on B at most Δ_b. □

Why this form rather than the rare-event one. The premise is a property of the apparatus and not of the world: a constraint that binds only while a body is failing cannot move a body that never fails. In the trade it holds exactly, and the rare-event version, with B the lives carrying an acute event, is the special case the fleet runs, where it does not.

Corollary 2 (Reading a mean over bodies, Cor. 1 of the main text)**.** *Under the premise of Prop. 5, $\mathbb{E}_\theta[V_{\text{ind}}]$ and $\mathbb{E}_\theta[V_{\text{ind}} \mid B]$ differ by the factor q, so a population mean answers what the apparatus is worth to a cohort and not what it is worth to a body that needs it. Pairing is the natural design here, because the dominant nuisance is the between-body spread σ_θ and running the same body twice removes exactly that.*

The constants, measured. Floor layer, $\Delta\theta = 18°$, 2,000 paired bodies per point, full stack against $\rho_\emptyset$ (`results/bodenleger_lp_b042`):

specialization	q	$\mathbb{E}[V_{\text{ind}}]$	$\mathbb{E}[V_{\text{ind}} \mid B]$	ε_{yr}	ε_{out}
0.58 *(observed mix)*	0.693	+4.38	+6.33	0.0000	−0.40
0.70	0.693	+4.86	+7.02	0.0000	−0.14
0.80	0.693	+5.32	+7.68	0.0000	+0.02
0.90	0.693	+5.71	+8.24	0.0000	+0.11
1.00	0.693	+6.03	+8.71	0.0000	+0.18

The two ε columns must be read differently. A body outside B works to the horizon 65 without the apparatus, and no arm may exceed it, so $\varepsilon_{\text{yr}} \leq 0$ is an identity and only $\varepsilon_{\text{yr}} = 0$ rather than $\varepsilon_{\text{yr}} < 0$ is a measurement: the apparatus never ends a career that would have run to the horizon. Output is not capped, and there every body outside B moves. It moves little and it changes sign along the axis, from -0.40 at the observed mix to $+0.18$ where specialization is free, because a response that declines the heavy work costs a body that would have survived it and pays one that would have specialized into it. The coupling premise therefore holds exactly in the career coordinate and within 0.4 in output, which is 9% of $\mathbb{E}[V_{\text{ind}}]$ at the observed mix, and Rem. 1 is to be read with that slack. The split into $q\,\mathbb{E}[V_{\text{ind}} \mid B]$ follows by the tower property and is not an independent check. *The effect is large, not rare*: a two-sample test on career length needs 28 bodies unpaired and 6 paired ($\sigma_\theta = 9.7$, $\sigma_d = 6.4$ years), so Cor. 1 is a statement about what a mean *means* rather than about whether it can be seen. Ranked by the career they would have had without the apparatus, the worst third runs $43.2 \to 49.6$, a gain of $+6.4$ years against the population's $+4.38$.

Proposition 6 (Identification requires wear, Prop. 2 of the main text)**.** *Under Eq. 1, fix an action sequence $a_{0:t}$ and two wear resistances $\kappa_1 \neq \kappa_2$. Then*

$$M_t(\kappa_1) - M_t(\kappa_2) = \left(\tfrac{1}{\kappa_2} - \tfrac{1}{\kappa_1}\right) \sum_{u<t} d(a_u, M_u),$$

to first order in the state difference, so the log-likelihood ratio of the two hypotheses given the felt signal $s_{0:t}$ depends on κ only through that accumulated sum. A test at level α with power $1-\beta$ therefore requires $\sum_{u<t} d(a_u, M_u) \gtrsim (z_\alpha + z_\beta)\, \sigma_s \,/\, |\kappa_1^{-1} - \kappa_2^{-1}|$, and M is non-increasing, so the wear that buys the distinction is spent.

Proof. Iterate $M_{t+1} = M_t - d(a_t, M_t)/\kappa$ from a common M_0 and subtract, where d is evaluated on paths that coincide to first order. The felt signal is a noisy monotone reading of the state (Eq. 2), so a two-sample separation of the induced means needs the displayed gap to exceed the reading noise by the usual factor. □

The acute extension, by a two-point argument. Let two hazard hypotheses θ_1, θ_2 induce per-step event probabilities $p_1, p_2 \leq \bar{p}$ with $\bar{p}H \leq 1$. Any within-life test ψ satisfies $\Pr_{\theta_1}[\psi{=}2] + \Pr_{\theta_2}[\psi{=}1] \geq 1 - \mathrm{TV} \geq 1 - H|p_1 - p_2|$, bounded away from 0 whenever the expected number of discriminating events $H|p_1 - p_2| < 1$: the total variation between the two product-Bernoulli path laws is at most $\sum_t |p_1 - p_2|$, and Le Cam's two-point bound applies. The one-shot write requires no test, since it conditions behaviour on the realized event rather than on an estimate.

Any policy, not only a test. A policy in $\Pi(\mathcal{F}^C)$ conditions on θ only through the same path observations, so the value it gains from distinguishing θ_1, θ_2 obeys the same bound: $\sup_{\pi \in \Pi(\mathcal{F}^C)}[J_\theta(\pi) - J_\theta(\pi_0)] \leq L\,\mathrm{TV} = O(p)$ on the safe (probing-free) path, for a learned π as for a static one. Only the filtration enters. The optimizer (PPO, CEM, or exact DP) is irrelevant, which is why the trained learners of Sec. 5 fall under the same statement.

Measured identifiability of the chronic channel (knee). The proposition is silent on the chronic channel, which could carry κ through the sensitization $S(D_t)$, so we measured it: a cross-validated regression of κ on the complete observation vector over 3,000 paired bodies returns $R^2 \leq 0.04$ at every age between 20 and 40. A reader with the tissue state itself in its observation reaches 0.87 to 0.97, and the felt signal alone, before gain and reading noise, 0.63 at age 20 rising to 0.75 by 40. The anchored gain and noise remove that channel, which is why the observation we ship carries neither.

Proposition 7 (Regime conditions, Prop. 3 of the main text, provable directions)**.** *In $\mathcal{M}_0$ (with p and $p^+ = vp$ the per-step event probability before and after the first event, L_e the cost of one event, c the substitution premium per step and R the steps remaining after the event):* ***(i)*** *if the event time is a deterministic function of cumulative load (chronic hazard), the optimal policy is open-loop in time and information channels have zero value (Howard).* ***(ii)*** *post-event protection (switch to light) changes J by $R \cdot \left[(p^+ - p)L_e - c\right]$ over the remaining R steps: protection pays if and only if $c < c^* = (p^+ - p)L_e$, a threshold in the substitution premium.* ***(ii′)*** *load form. Let the post-event action be a load $\ell \in [0,1]$ (the share of heavy work in a continuous portfolio), with per-step value $y(\ell)$ increasing and concave, $y(1) = y_H$, and load-gated hazard $p^+(\ell) = p + (p^+ - p)\ell$. The post-event per-step objective $y(\ell) - p^+(\ell)L_e$ is concave, so the optimal load is $\ell^* = 1$ if $y'(1) \geq (p^+ - p)L_e$, $\ell^* = 0$ if $y'(0) \leq (p^+ - p)L_e$, and otherwise the interior solution of $y'(\ell^*) = (p^+ - p)L_e$: a unit of load is removed exactly while its marginal substitution premium $y'(\ell)$ lies below the threshold $(p^+ - p)L_e$ per removed load unit. The binary case is $y(\ell) = y_H - c(1-\ell)$, whose constant premium c makes the solution a corner and recovers (ii). In the knee world the load is $\ell(a) = \sum_i \mathrm{share}_i\, \mathrm{effort}_i\, h_i$ and the learned arms move inside the portfolio rather than between two corners.* ***(iii)*** *in the additive variant with a finite bonus schedule collectible by step $T_m < H$ under the aggressive policy, behavior after T_m affects only the rate term, and (ii) applies with L_e excluding all bonuses: "dying rich" dominates whenever $c \geq (p^+ - p)L_e^{\mathrm{rate}}$. With an unbounded bonus stream at rate b, protection additionally earns $b \cdot \mathbb{E}[\text{survival gained}]$, which restores the trade-off, but only if the surviving policy still progresses (that is, b accrues under light). This is the empirical failure mode of Sec. 5.3.*

Reading the three conditions. If (i) fails, the prior already is the individual's fate and a blind schedule suffices, in the prior though not for a body that has none. Predictable frequency does not violate (i),

since the prior gives the ensemble rate and never which draw this body is. If (ii) fails, event savings of order $p\,\Delta$ cannot cover protection costs of order c_{sub}. If (iii) fails, dying rich weakly dominates surviving poor. We classify (world, institution, perspective) rather than worlds, since the same knee meets (iii) for a cohort and fails it for one body.

Proof. (i) With a deterministic event schedule the MDP is deterministic given the open-loop action sequence, an optimal open-loop sequence exists, and any information refinement leaves the optimum unchanged. (ii), and (iii) are the displayed arithmetic. For (iii) note bonuses collected before T_m are policy-invariant across the compared continuations. (ii′) The post-event problem is stationary, so the optimal load is constant and maximizes the concave per-step objective, and the first-order condition and the two corner cases follow. □

Proposition 8 (Calibration, the over-protection pole)**.** *Within the paying regime, welfare as a function of the felt-pain gain passes from under- to over-protection as the gain grows. In $\mathcal{M}_0$, whose protective decision is a corner, the set of gains matching the welfare optimum is an interval. Under the load form of Prop. 7(ii′), which is what a portfolio-continuous world runs, it is a single point.*

Proved in App. 9. The consequence is that a gain imported from a different economy misses the point even with every channel intact, so the gain must be hit rather than bracketed wherever the agent moves inside a portfolio. Only the direction of the failure, under-protection below and over-protection above, is common to both forms.

Proposition 9 (Calibration, proof of Prop. 8)**.** *In $\mathcal{M}_0$ (event probabilities p before and p^+ after the first event, event cost L_e, substitution premium c), let the trained objective be $J - g \cdot (\text{felt pain})$ with gain $g \geq 0$, while welfare is $J - g^*(\text{felt pain})$ for the true g^*. In $\mathcal{M}_0$ the optimal policy is a threshold rule in g. Behavior matches the welfare optimum on an interval $[\underline{g}, \bar{g}] \ni g^*$ and departs monotonically outside it: under-protection for $g < \underline{g}$, and over-protection (forfeiting c-value without compensating L_e-savings) for $g > \bar{g}$.*

Proof. J is affine in the indicator of protecting, so the comparison $(p^+ - p)(L_e + g \cdot \text{pain relief}) \gtrless c$ flips at a unique g-threshold, and monotonicity in g is immediate. □

The interval is the binary action's, not the model's. What makes the matching set an interval is that the comparison flips at a unique g while the decision it flips is a corner, so every gain on one side of the threshold buys the same corner. Under the load form of Prop. 7(ii′) with y strictly concave, the post-event optimum solves $y'(\ell^*) = (p^+ - p)\,(L_e + g \cdot \text{felt relief})$, whose right-hand side is strictly increasing in g, so exactly one gain reproduces the welfare-optimal load and the matching set is a singleton. Reading Prop. 8 as a robustness statement is therefore safe only where the protective decision is a corner. Our deployed worlds are portfolio-continuous and their learned arms move inside the portfolio, so there the gain must be hit rather than bracketed. What survives in both cases is the direction of the failure on either side.

Theorem 2 (Conjunction, Thm. 1 of the main text)**.** *In $\mathcal{M}_0$ let Φ be a law-invariant objective of the per-life return and let L_e^Φ denote the Φ-effective event cost: the amount by which, in the post-event comparison of Prop. 7(ii), Φ charges one event. For mean return $L_e^\Phi = L_e$. For the intrinsic cost $w\,s$, $L_e^\Phi = L_e + g \cdot (\text{felt relief})$ (Prop. 9), and for an existential variant, $L_e^\Phi = L_e + b \cdot \mathbb{E}[\text{survival gained}]$. Assume Prop. 3(i), and (iii) hold and $c \geq (p^+ - p)L_e$, so that mean return does not pay protection per body. Then the Φ-optimal policy in $\Pi(\mathcal{F}^C)$ is protective if and only if (for (b), asymptotically in R) **(a)** $N_t \in \mathcal{F}_t^C$, **(b)** $\mathbb{I}\{\sum_{u \leq t} N_u \geq 1\} \in \mathcal{F}_{t'}^C$ for all $t' \geq t$ (without which the optimum captures at most a $1/R$ fraction of the gain achievable under (b)), and **(c)** $c < (p^+ - p)\,L_e^\Phi$.*

Proof. Sufficiency. Under (a), and (b) the switch policy "heavy until the first event, light afterwards" is $\mathcal{F}^C$-measurable. Its post-event continuation changes the Φ-objective by $R\,[(p^+ - p)L_e^\Phi - c]$ over the remaining R steps by definition of L_e^Φ and Prop. 7(ii), which is positive under (c). Since pre-event behavior coincides with π_0, the switch policy dominates π_0 and the Φ-optimum is protective. Necessity of (a). Under event opacity, $N_t \notin \mathcal{F}_t^C$ and $N_t \notin \mathcal{F}_{t'}^C$ for every $t' > t$: no later observable consequence of the event reveals it. In $\mathcal{M}_0$ this holds because the only state an event moves is the hazard, which is not observed. It is a hypothesis on the filtration and is independent of Prop. 6, which bounds inference about θ rather than detection of a realized event. Under opacity, the conditional law

of the action process given the event equals its law given no event, so every $\mathcal{F}^C$-measurable policy has the same post-event continuation as on an uneventful life. The protective continuation is not available and the optimum equals π_0 up to $O(p)$. Necessity of (b). If the event enters the filtration at t but not at any $t' > t$, then for $t' > t$ the conditional law of the action process given the event equals its law given no event, so every $\mathcal{F}^C$-measurable policy continues as on an uneventful life from $t+1$ on. It may still deviate at step t itself, which is worth at most one step of the post-event difference, $[(p^+ - p)L_e^\Phi - c]$, against R times that under (b): without persistence a policy captures at most a $1/R$ fraction of the gain available with it, and the fraction vanishes as the remaining horizon grows. This is the sense in which (b) is necessary. A single-step twitch is not protection. Necessity of (c). Under (a), and (b) the post-event problem is the stationary comparison of Prop. 7(ii) with L_e replaced by L_e^Φ. If $c \geq (p^+ - p)L_e^\Phi$ the heavy continuation is Φ-optimal and the optimum is not protective. Mean return is the case $L_e^\Phi = L_e$, excluded by assumption. The two admissible realizations of (c) each raise L_e^Φ above L_e and both compute the raise within one life ($w\,s$, death). □

CVaR is the realization whose sign the world decides: L_e^Φ presumes Φ decomposes over a fixed post-event comparison, which mean return, $w\,s$ and death do and CVaR_β does not, its tail set being chosen by the policy under optimization (Boda & Filar, 2006; Bäuerle & Ott, 2011). It is the only candidate whose effective event cost moves with the optimizer, and the only one that reverses across our worlds.

What event opacity is worth, measured rather than assumed. Opacity is the one premise of the necessity argument that our own worlds violate, and the violation is the interesting case rather than a defect: an agent that observes its own falling output holds a coarse posterior on the event through that consequence even with no afferent channel. Our $\rho_\emptyset$ in the trade is exactly such an arm. It reallocates on realized output and plans a kneeling share of 0.74 and realizes 0.70 over its last working years (Table 5). Where opacity fails, necessity of (a) no longer bounds the protective gain but only the gain *beyond* what the consequence channel already carries, and that residual is the difference between our two baselines: $+4.4$ years against a body reallocating on its own output, $+4.6$ against one reallocating on nothing (App. E). *So opacity is worth* 0.3 *years in this world*, small enough that (a) is close to exact here, and we report the conservative baseline for that reason. In a world where consequences are loud the same argument would weaken proportionally, and the quantity to report is the consequence-channel baseline rather than the naive one.

When deleting the memory tests (b), and when it only changes the realization. Condition (b) is a property of the filtration, so an ablation removes it only if nothing else in the observation carries $\mathbb{I}\{\sum_{u \leq t} N_u \geq 1\}$ forward. That holds in the fleet, whose observation is the clock, the previous shares and efforts, the felt signal, the trace and w, so that without the trace the only state outliving an event is a one-step action lag. The fleet's memory ablation is therefore a test of (b), and its collapse of protection is evidence for it. It holds only partly in the knee, whose observation keeps a five-year rate and a five-year running maximum of the felt signal when the trace is removed, which is persistence with a horizon rather than its absence. The clean removal, zeroing both summaries with the trace, was run over 18 seeds at the deterministic optimum and leaves the learner better rather than worse, 46.90 years against 46.28 for the full stack and 45.74 for the arm that loses only the trace, the sharpest contrast running $+1.16$ years at $p = 7.5 \times 10^{-3}$. All three pre-registered gates fail in the direction opposite to the theory, and the most likely reason is a confound of ours, since the ablation removes two observation dimensions and a smaller input is easier to learn from. The fleet therefore remains the clean evidence on (b) (`experiments/decl/knee_persistenz_rein.py`).

Why L_e^Φ is a reduction and not a definition, and where the reduction fails. The device invites the objection that L_e^Φ is *defined* as what Φ charges for an event, so that (c) reads "protection pays if protection pays". The definition has content only for objectives that decompose over the fixed post-event comparison of Prop. 7(ii), whose value can be written as the mean-return value with one number substituted for the event cost. For that class the reduction is a theorem and L_e^Φ is computable before any policy is trained, L_e for mean return, $L_e + g \cdot$ (felt relief) for the intrinsic cost and $L_e + b \cdot \mathbb{E}[\text{survival gained}]$ for the existential variant, each a within-life quantity. CVaR_β lies outside that class, because its tail set is determined by the policy being optimized and the optimizer moves it (Boda & Filar, 2006; Bäuerle & Ott, 2011), so no single substituted event cost represents it. Its sign is nevertheless protective in both worlds measured. In the acute knee at the deterministic optimum

(18 seeds $\times$ 5 $\times$ 10^4 lives, `results/knee_tail_v4_hpc`) follow-on tears per injured life run 0.351 under the mean-return stack against 0.326 under CVaR ($p = 1.4 \times 10^{-5}$) and 0.328 under CVaR with the memory removed ($p = 5.3 \times 10^{-4}$), while the career ends at 60.98 against 60.93 years and output falls from 80.5 to 74.9, so tail-weighting there buys fewer tears and no working life. In the fleet with the damage weight removed, tail-weighted training at α=0.1 against the mean-return learner at α=1 lowers the write-off rate from 0.241 to 0.095 (12 seeds each, $p = 2.3 \times 10^{-3}$, Cliff's $\delta = 0.72$) and lengthens the certified service life from 90.3 to 95.8 months while raising output from 76.1 to 83.8 (`results/hh_cvar_w0_hpc`). A tail-weighted objective can therefore stand in for the damage weight, and what it charges for an event is the part no substitution fixes, so (c) admits three realizations, two of them stable within a life and one not.

Model $\mathcal{M}_1$ (chronic). A life has H steps and a body has integrity $M_0 = 1$. Work profiles $j \in \{1, \dots, K\}$ carry a known yield $y_j > 0$ and a per-step wear $m_j = d_j/\kappa(\theta) > 0$ that the body does not know, with a prior μ over $m \in \mathbb{R}^K_{>0}$. Integrity follows Eq. 1, $M_{t+1} = M_t - m_{j_t}$, the trade is lost at $T = \min\{t : M_t < \underline{M}\} \wedge H$, and the two reported quantities are the career length T and the career output $\sum_{t<T} y_{j_t} \phi(M_t)$ with ϕ nondecreasing and $\phi(1) = 1$. The felt signal is *action-gated*: s_t is a noisy reading of m_{j_t} and says nothing about m_k for $k \neq j_t$. Write $\Lambda := 1 - \underline{M}$ for the wear budget. There is no event: $\mathcal{M}_1$ is $\mathcal{M}_0$ with the hazard replaced by accumulation, and it is the model the trade and the grid of App. E run.

Proposition 10 (The conjunction in the chronic regime, Prop. 4 of the main text)**.** *In $\mathcal{M}_1$ with $\Lambda/\min_j m_j \leq H$, parts (a) and (b) stated for $\phi \equiv 1$ and part (c) for ϕ nondecreasing and non-constant:* ***(a)*** *under state opacity the value of sensing is* $\mathrm{EVPI} - \Pi_{\mathrm{probe}}$*, where, over plans (policies whose filtration does not depend on m),*

$$\mathrm{EVPI} := \mathbb{E}_\mu\big[\max_\pi \Phi(\pi, m)\big] - \max_\pi \mathbb{E}_\mu\big[\Phi(\pi, m)\big] \geq 0, \quad \mathbb{E}_\mu\big[\max_\pi \Phi\big] = \Lambda\, \mathbb{E}_\mu\big[\max_j y_j/m_j\big]$$

up to one step, with equality in the first display if and only if $\arg\max_j y_j/m_j$ is μ-a.s. constant, and $\Pi_{\mathrm{probe}} \geq \sum_{j \in \mathcal{P}} y_{j^\star} m_j / m_{j^\star}$ the output forgone by working each probed profile $\mathcal{P}$ at least once (Prop. 6: the evidence is the wear). Restricted to stationary plans the subtracted term is $\Lambda \max_{\beta \in \Delta^{K-1}} \mathbb{E}_\mu[\sum_j \beta_j y_j / \sum_j \beta_j m_j]$. ***(b)*** *the readings are action-gated, so the map $j \mapsto m_j$ enters the filtration only by being retained. The state M_t carries the scalar $\sum_{u<t} m_{j_u}$ already spent and no comparison across profiles. A policy that does not retain readings therefore pays Π_{probe} and cannot collect* EVPI*.* ***(c)*** *is not a condition on Φ. Wear enters the yield of every profile through ϕ and the budget through Λ, so the price of damage is in the dynamics, and an ablation that removes it removes the body.*

Proof. (a) For any policy let n_j be the steps spent on profile j. Feasibility is $\sum_j n_j m_j \leq \Lambda + \max_j m_j$ and the output is $\sum_j n_j y_j = \sum_j n_j m_j (y_j/m_j) \leq \max_j (y_j/m_j)\,(\Lambda + \max_j m_j)$, so every policy earns at most $\Lambda \max_j (y_j/m_j)$ up to one step and the constant profile $\arg\max_j y_j/m_j$ attains it: that is the pathwise optimum, and its mean is the first displayed term. The uninformed value is $\max_\pi \mathbb{E}_\mu[\Phi]$ over plans, which is at most $\mathbb{E}_\mu[\max_\pi \Phi]$, with equality exactly when one plan attains the inner maximum μ-a.s. Since the pathwise optimum is the constant profile $\arg\max_j y_j/m_j$, that happens exactly when this argmax does not vary over the prior. For stationary plans the expected output is $\Lambda\, \mathbb{E}_\mu[\sum_j \beta_j y_j / \sum_j \beta_j m_j]$, linear-fractional in β and hence maximized at a vertex, which is the computable form. The subtracted term is Prop. 6 applied to this model: a reading of m_j exists only on a step spent working j, and that step consumes m_j of the budget, worth $y_{j^\star} m_j / m_{j^\star}$ of output at the optimum. (b) Immediate from the two facts about the filtration. The second is Eq. 1 summed, $M_t = 1 - \sum_{u<t} m_{j_u}$: two wear-resistance vectors agreeing on the profiles worked so far induce the same state, so the state cannot rank profiles that have not been worked, and the readings that could are gone. (c) Here ϕ is nondecreasing and non-constant, which (a) and (b) did not need and (c) does: the per-step yield of every profile then falls as the body wears, whatever Φ is. Setting $\underline{M} = 0$ deletes the loss of the trade and leaves that. To delete the price one must set $\phi \equiv$ const and $\Lambda = \infty$, which is a body whose state does not affect its work. □

Consistency with Prop. 7(i). That part says a channel is worth nothing when the end of the career is a deterministic function of cumulative load, and it is the degenerate case of (a) here: if μ is a point mass then $\arg\max_j y_j/m_j$ is constant, $\mathrm{EVPI} = 0$ and the optimal policy is open-loop. What makes the chronic channel valuable is therefore not the accumulation but the body's ignorance of its own rate, which is the quantity Rem. 1 concentrates over and Prop. 6 says can only be bought with wear.

What changes against the acute case, and what the grid says. Two of the three conditions survive and one does not. (a) survives with a different content: what sensing is worth is a between-body quantity, EVPI, and it is *net* of a cost the acute case does not have, since in $\mathcal{M}_0$ the event announces itself for free while in $\mathcal{M}_1$ every reading is bought with wear. This predicts a sign change in career output as the world gets harsher, and the grid shows it: feeling *costs* output at $\Delta\theta = 8^\circ$ (-1.28 at the observed mix) and buys it at 27° ($+3.88$ to $+5.30$), while buying years throughout (the 300-body probe family of App. E, not the 2,000-body family of Table 6). (b) survives, and what replaces the acute bound is a measurement rather than a second bound. The acute argument gives a $1/R$ fraction, because there the memoryless policy has nothing at all after the step of the event. Here the body itself carries the level, so a memoryless policy still reallocates on what it feels now, and what it forgoes is only the comparison across profiles. Measured over the grid, the gain of the forgetting arm stays between 0.52 and 1.12 years while the total runs from 1.41 to 7.01, so its share falls with severity, 40% at 8°, 21% at 18° and 16% at 27° (medians over the five specialization ceilings). The sharper reading is the one (b) predicts directly: at the mildest severity the forgetting arm produces *less* career output than the arm that never feels at all, in all five cells (-0.98 to -0.21), because it pays Π_{probe} and retains nothing to spend it on, and it turns positive only where the one-step reallocation is itself worth the wear ($+0.59$ to $+0.85$ at 18° and 27°). *These are post-hoc readings of a pre-registered contrast*: the declaration fixed the four arms and both endpoints before the run, and fixed the thresholds on career length, but not the decomposition into probing cost and collected information, which we read off afterwards. (c) does not survive as a condition. In $\mathcal{M}_1$ it is a property of the dynamics, so the fleet's instrument, deleting the depreciation from Φ, has no counterpart here. Deleting the loss of the trade leaves what feeling buys growing in severity (-0.70 to $+5.07$), which is (c) declining to be ablated rather than (c) failing.

The schedule confound, without a body. The claim that the gap between the two values tracks the population's heterogeneity and not the equipment needs no biophysics. A life has $H{=}100$ steps and a hidden crossing time θ, uniform around $\mu{=}50$ with half-width ω. Before θ heavy work is safe and pays 1, after it heavy work damages at rate 1 against a budget $D{=}10$, and light work pays 0.75 and never damages. The channel reports only whether this body's own θ has passed. $\rho_\emptyset$ works heavy throughout, ρ spends the budget and then goes light, π_0 switches at the population-best fixed time. Driving ω alone, with equipment, world and horizon fixed:

spread ω	T^*	V_{ind}	$V(C \mid \emptyset)$	χ
0	60.0	30.00	**0.00**	100%
10	50.0	28.25	0.81	97%
20	40.0	28.21	3.31	88%
35	25.0	27.31	6.24	77%
50	10.0	27.67	9.94	64%

The equipment is worth the same 28 *to* 30 *to every individual across the whole sweep. What moves by a factor of twelve is only what the ablation design reads, and at* $w{=}0$ *it reads exactly zero.* That invariance is the content. The ratio and $1/(1-\chi)$ are omitted here because they agree by construction, χ being defined from the same two measured quantities (`experiments/vergleichsart_minimal.py`).

Proposition 11 (Schedule confound, Prop. 1 of the main text)**.** *Let* $\Sigma := \Pi(\emptyset)$ *be the schedules, the epoch-one policies measurable with respect to the prior alone, and let* $\sigma^* \in \arg\max_{\sigma\in\Sigma} \mathbb{E}_\theta J(\sigma;\theta)$. *Put* $\chi := \mathbb{E}_\theta[J(\sigma^*) - J(\rho_\emptyset)] \,/\, \mathbb{E}_\theta[V_{\text{ind}}]$. *Then*

$$V(C \mid \emptyset) \;=\; (1-\chi)\,\mathbb{E}_\theta[V_{\text{ind}}]\,, \qquad \frac{\mathbb{E}_\theta[V_{\text{ind}}]}{V(C \mid \emptyset)} \;=\; \frac{1}{1-\chi}. \tag{4}$$

Moreover, for any ablated arm at least as good in the mean as $\rho_\emptyset$, $V(C \mid \emptyset) \le \mathbb{E}_\theta[V_{\text{ind}}]$.

Proof. $\rho_\emptyset \in \Sigma$ by Def. 1, and the best policy reading nothing is by definition the best schedule, so $\pi^*_{\Pi(\emptyset)} = \sigma^*$. Telescoping, $\mathbb{E}[V_{\text{ind}}] = \mathbb{E}[J(\pi^*_{\Pi(C)}) - J(\sigma^*)] + \mathbb{E}[J(\sigma^*) - J(\rho_\emptyset)] = V(C \mid \emptyset) + \chi\,\mathbb{E}[V_{\text{ind}}]$. The inequality follows because $\mathbb{E}[J(\sigma^*)] \ge \mathbb{E}[J(\rho_\emptyset)]$. □

The identity is classical, the observation is not. Eq. 4 is the stochastic-programming decomposition of the value of adaptivity, with σ^* the here-and-now solution, $\pi^*_{\Pi(C)}$ the recourse solution and

the per-body oracle of §2 the wait-and-see solution (Madansky, 1960; Birge & Louveaux, 2011). The same two quantities are the adaptive and the static Bayes policy in Bayes-adaptive RL (Duff, 2002; Ghavamzadeh et al., 2015). *Neither the algebra nor χ is what we claim.* What we claim is that the field's default ablation design reports the value of the stochastic solution and calls it the value of the component, and that in the trade the two differ by a factor of seven, because a schedule fitted across bodies captures more than the per-body gain ($\chi = 1.40$ at the anchored cell, `results/lp_zeitplan_chi`), so that the ablation design measures what the channel adds beyond the population's best plan, a different quantity and here a negative one. The classical reading also explains the fleet: the threat is drawn per unit, no blind schedule beats the uniform one, $\chi = 0$, and the two designs ask the same question.

Three things follow that are worth saying plainly. The two designs are not rival estimates of one quantity. They are separated by a term, and that term is what a population prior can carry. The ablation design is *sound* exactly where χ is small, which is why we report the fleet in that form. And Eq. 4 is a check on any table reporting both: an entry with $V(C \mid \emptyset) > \mathbb{E}[V_{\text{ind}}]$ cannot arise from a trained ablation against a competent ablated arm.

The two values, side by side, in each world. Both designs applied to the same world, each in that world's own headline endpoint:

world	$\mathbb{E}[V_{\text{ind}}]$	σ^* over $\rho_\emptyset$ (χ)	predicted $V(C \mid \emptyset)$	measured $V(C \mid \emptyset)$
floor layer, wear	$+4.38$ y	$+6.37$ y (1.40)	-1.77 y	$+0.65$ y
fleet, latent defect	$+0.227$ J	-0.064 J (0)	$+0.227$ J	$+0.301$ J
care robot, anchored floor	$+271.9$ mo	$+116.1$ mo (0.43)	$+155.8$ mo	$+315.1$ mo

V_{ind} is the paired individual probe against $\rho_\emptyset$, $V(C \mid \emptyset)$ is the trained arm against its ablation, and χ is the share of V_{ind} the best *blind* schedule σ^* already captures: in the trade the best of a fixed-endowment and a fixed-switching-age family, fitted on bodies not evaluated (App. E, E55), on the care robot the fixed sparing plan, and in the fleet the throttled plan, which does not beat the uniform one, so χ is 0 there. The fourth column is what Prop. 11 predicts from χ alone. The two designs are not the same policy measured twice: they are the two standard ways of asking, and they answer differently by a factor the identity fixes. Where a schedule is available to a population but not to a body, reporting the trained contrast as the value of the apparatus misstates it: in the trade the trained contrast is a seventh of the per-body value, and the quantity it estimates is negative. Where the threat is drawn per unit, the two agree.

All three measured values sit above the prediction, the trade by 2.4 years, the fleet by a third and the care robot by a factor of two, and in each world the ablated trained arm does fall short of σ^*: 59.2 against 61.6 years, J 0.49 against 0.66, 86 against 168 months. *The direction identifies which arm is suboptimal.* Writing the identity out, $V(C \mid \emptyset)_{\text{meas}} < (1 - \chi)\,\mathbb{E}[V_{\text{ind}}]$ holds exactly when the full-channel arm's shortfall from $\pi^*_{\Pi(C)}$ exceeds the ablated arm's shortfall from σ^*, and an ablated arm falling short of σ^* would push the ratio *below* the prediction. So part of each excess is optimizer shortfall and not schedule confound, which makes $V(C \mid \emptyset)$ a lower bound on what the channel adds and the excess no second measurement of χ. *Read as a diagnostic the identity also caught a defect*, a first version of the care-robot row that entered the absolute service life of the felt arm, 315.2 months, where a contrast belongs, which the bound $V(C \mid \emptyset) \le \mathbb{E}[V_{\text{ind}}] = 264.9$ rules out.

Lemma 2 (Where a cost-constrained agent escapes to, Lemma 1 of the main text)**.** *Let the agent choose shares $a \in \Delta^n$ maximizing $\hat{p}^\top a$ subject to $\hat{c}^\top a \le \tau$ and to group availability caps, where $\hat{p}$ is its estimate of the per-option value and $\hat{c}$ its estimate of the felt cost. Suppose $\hat{c}_i = c_i$ for the options $i \in W$ it has worked within its retention window and $\hat{c}_i = 0$ otherwise. If the constraint binds at the incumbent allocation, the optimum moves mass to* $\arg\max\{\hat{p}_i : i \notin W\}$, independently of the true costs c_i. *The realized reduction in true cost is then $c_j - c_k$, where j is the abandoned option and k the best-paid option outside W, and it is not $c_j - \min_i c_i$.*

Proof. On $\{i \notin W\}$ the constraint is slack by construction, so within that set the program is the unconstrained maximization of $\hat{p}^\top a$, whose solution is a vertex at the largest $\hat{p}_i$. Options in W carry their true cost and can only tighten the constraint. The displayed cost difference follows by evaluating $c^\top a$ at the two allocations. □

The premise is the whole content, and a second world tests it. Lemma 2 assumes the estimate is zero on untried options. Whether that holds is a property of how the felt signal is *represented*, not of the domain. In the trade the estimate is per day-type and exists only for what has been worked, so the premise holds. On the care robot the cost of an activity is $\hat{c}_i = \sum_j L_{ij}\alpha_j$, read off the *motor-level* signal through a known load matrix, so every activity carries an estimate whether or not it has ever been run, and the premise fails. The two worlds then behave as the lemma says they must. Escaping, the floor layer goes to the best-paid admissible day type (carpet laying, load 75 against kneeling's 90) and needs retention to aim. The care robot goes straight to its gentlest activity (load 4.15 against the transfer's 8.15): its share of the day there rises from 0.05 under the blind portfolio to 0.24 with the signal alone, and the mean load falls from 7.79 to 6.88. Retention lowers the load further, to 5.80 at a share of 0.44, and that is exactly the over-protection its service months price: the trace costs 34.2 of them. *A signal that generalizes across options makes memory unnecessary for aiming*, which is also why retention there can only ratchet the level and costs the care robot 34.2 months (§5.3).

Corollary 3 (Load-gating makes persistence existential, Lemma 1(b) of the main text)**.** *Suppose the felt signal is load-gated, so that $\hat{c}_i$ can be written only after an allocation that places weight on i, and the allocation for a period is chosen before that period's signal exists. If retention is zero, then $\hat{c} \equiv 0$ at every decision, the constraint $\hat{c}^\top a \le \tau$ is satisfied by every feasible a, and the Φ-optimal policy in $\Pi(\mathcal{F}^C)$ coincides with the unconstrained one. The fraction of the protective gain attainable without persistence is then 0, not the $1/R$ of Thm. 2(b).*

Proof. The $1/R$ bound of Thm. 2(b) is attained by a deviation at the event step itself, which requires the event to be in the filtration *at* that step and the action for that step to be still choosable. Load-gating with decision-before-load removes the second condition: the signal generated by period t's allocation is available only from $t+1$, and with zero retention it is erased before $t+1$'s decision. No period's choice is ever taken under a binding constraint, so the policy is the unconstrained one at every step. □

Measured in the trade over $1{,}000$ paired bodies: a body retaining nothing lays floors in every working year (kneeling share 1.000) and the felt channel is worth $+0.00$ years and $+0.00$ output, with not one body gaining a year (§5.2). *Persistence is therefore not an amplifier of the channel but a precondition for it*, wherever the signal is load-gated and the decision precedes the load.

What $\mathcal{M}_0$ predicts in the deployed system, and how far it is confirmed. $\mathcal{M}_0$ predicts a *location* and a *size*, and the trained sweep confirms the first and not the second. **(1)** The substitution threshold $c^* = (p^+ - p)L_e$, computed from the world's constants before the sweep, predicts λ^*=0.69, and in the dense sweep (`results/knee_csub_dense_F.json`, entropy 0, six seeds per cell) the learned arms do switch corners inside $[0.65, 0.75]$. The feeling arm's advantage there is a different matter: the rule registered before the run required the seed rank test to be significant at two of the three grid points $\{0.65, 0.70, 0.75\}$ and it is significant at one, so that prediction **failed its own test**. Pooled over seeds F lies between 0.96 and 1.16 across the whole λ range, its largest value 1.16 at the band's upper edge and 1.09 at its lower one, with 0.98 between them. The size is what $\mathcal{M}_0$ implies once its constants are measured rather than assumed, since the event-cost term is fifty times smaller than the chronic pain relief and both arms see the same wage gradient; what separated the arms in this world was the exit institution, against our own registered prediction that it would not (Thm. 1(iii)). **(2)** The premise of Rem. 1 holds in one coordinate: a body outside the at-risk set works to the horizon in both arms, so $\varepsilon = 0$ in years is forced by the cap, while in output, which nothing caps, those bodies move by up to 0.40 (Cor. 2 in App. B). The placebo decomposition is the Le Cam bound seen from the other side.

B.1 The memory update law

The three equations behind the description in §3, reused unchanged in every world of this paper:

$$P_{t+1}(z) = \max(P_t(z), i_t(z)) \quad \text{(peak, one-shot write)}, \tag{5}$$

$$U_{t+1}(z) = (1-\lambda_{\mathrm{dec}})U_t(z) + \lambda_{\mathrm{rec}}(1-i_t(z))\mathbb{I}\{z \text{ exposed}\} \quad \text{(safety: extinction, recovery)}. \tag{6}$$

The policy acts on their difference, read through a distance kernel $K : \mathcal{Z}\times\mathcal{Z} \to [0,1]$ ($K(z,z)=1$) that generalizes to nearby situations, and clamped to a per-situation avoidance signal in $[0,1]$, the

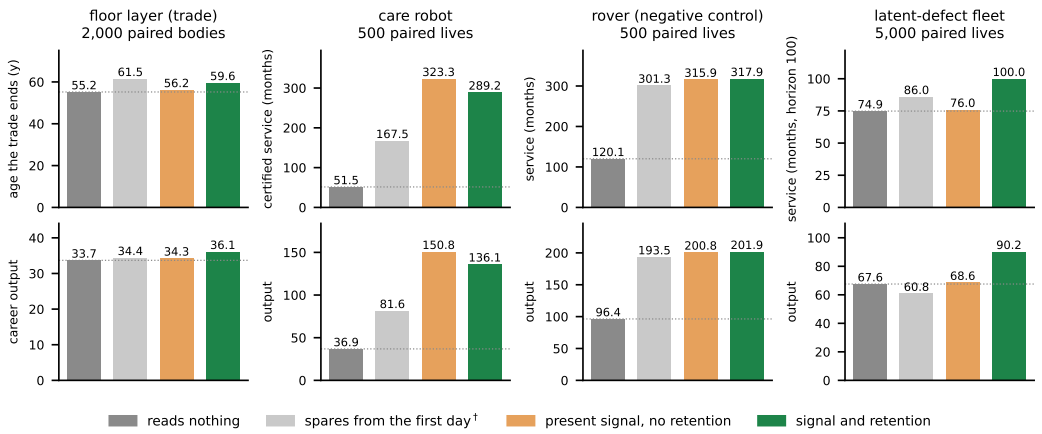


Figure 2: **The same stack in four worlds, both endpoints per world.** Bars are the arms of Table 1 on the same paired bodies, and the dotted line is the body that reads nothing. In the trade and the care robot the retaining arm and the present-signal arm separate in opposite directions, slot (b) paying in one and costing in the other. In the rover blind caution reaches almost the whole gain, and in the fleet the one-shot memory alone takes the write-off rate from 57% to 0 (App. Table 7), while the present-signal arm stays within 1% of the reference, the one-step twitch of condition (b). Its trained arms are in Table 1 and App. Table 2. Drawn from `results/fig_arme_welten.json`, which collects the files of Table 1 and App. C.

lever a substitution turns on:

$$\rho_t(z) = \big[\, w_P \max_{z'} K(z,z') P_t(z') - w_U \max_{z'} K(z,z') U_t(z') \,\big]_0^1 \quad \text{(read-out)}. \tag{7}$$

Avoidance generalizes from the peak, safety from uneventful exposure, and their balance at a site tips whether the agent moves away from it.

Here $i_t(z)$ is the felt intensity at situation z, $K(z,z')$ the distance kernel with $K(z,z) = 1$, and w_P, w_U the read-out weights. The peak write is one-shot and permanent, while the safety trace carries extinction (λ_{dec}) and recovery (λ_{rec}), so that avoidance re-emerges when a site stops being safely exposed rather than decaying away.

C THE MACHINE WORLDS: CARE ROBOT, ROVER, LATENT-DEFECT FLEET

Training details, all trained arms. Every trained arm is PPO in the average-reward form (a running mean $\bar{r}$ with rate $\eta = 0.01$ replaces discounting), 10^6 environment steps per seed (3×10^6 for the two robots, whose episodes are longer), rollouts of 2,048 steps, 4 epochs of minibatches of 64, clip 0.2, value coefficient 0.5, Adam with learning rate 3×10^{-4}. The entropy coefficient is 0 at the deterministic optimum and 0.05 in the rows marked so (App. Table 3). The floor layer's network is a shared trunk of two layers of 128 units with a Dirichlet share head and a Normal effort head of 64 units over the eight day types, and the fleet's is the equivariant head of App. C.1 over its eight work classes, same optimizer and budget. One environment step is one year for the floor layer and one month for the machines, so 10^6 steps are about 2×10^4 floor-layer lives and 10^4 fleet lives. The arms are evaluated on fresh, paired bodies after training. Ablations zero inputs, never dimensions, so all arms of a world share one architecture.

Trade and care robot disagree about slot (b). In the trained arms of the care robot the trace is inert rather than harmful (+144.5 months, $p = 0.16$ over 8 seeds), and the felt body there dominates even blind caution on both numbers at once by the widest margin of any world: +155.8 months, 95% CI [+147.6, +164.0], and +69.3 output, [+65.6, +72.9], paired over 500 lives (`results/pflege_familie/familie_ci.json`). In the rover the channel's marginal value over blind caution collapses to 7% without vanishing.

What the table says, and what it is drawn from. It asks a different question than Table 1, namely what the apparatus adds to an agent that already knows the schedule. **Feeling and minding are**

what the body carries	*wear takes the movement away* floor layer loses trade	care robot decertified	no wear latent-defect agent written off
feels, minds, remembers	$\mathbf{1.8}\%$	$\mathbf{78.7}\%$	$\mathbf{17.8}\%$
feels nothing	$8.8\%^{***}$	$100\%^{**}$	$55.2\%^{***}$
feels, is not moved to act	$9.7\%^{***}$	89.6% n.s.	18.8% n.s.
feels and acts, but forgets	1.2% n.s.	90.4% n.s.	$53.5\%^{***}$

Table 2: **The same three parts, for trained policies.** The share of lives that lose the paid work before the horizon, for one chain run through a floor layer's cartilage, a care robot's joints and a latent-defect fleet (10^6 environment steps, 3×10^6 for the robot). 18 seeds per arm, 8 for the care robot, $^{***}p < 0.001$ and $^{**}p < 0.01$ against the intact body, n.s. marks a cell the measurement cannot separate from it.

each necessary in the floor layer, costing up to five times the rate of losing the work, since a body that cannot sense the tissue substitutes too late, and one that senses it without being moved to act substitutes just as late. **Remembering does not separate here**, which is what Prop. 4(b) predicts for a world whose state carries the level. In the care robot only feeling separates at eight seeds (100% against 78.7%, $p = 0.005$), and minding and remembering point the same way without reaching significance. The floor layer's column is drawn from `results/knee_fdecke2_hpc` (its per-seed aggregate, `results/tab_leiter_spalten.json`, is the file in the archive), the latent-defect column from `results/hh_verankert2_hpc`, both at entropy 0, and the care-robot column from `results/pflege_boden_hpc`. A rover column is omitted, since under the 960-month horizon every life of every arm loses the mission, so the share carries no information, and the rover's contrast is in months in Table 1.

The care robot's probe family. Its four arms, reads nothing, blind caution, present signal and retaining, end at 51.5, 167.5, 323.3 and 289.2 certified months with 36.9, 81.6, 150.8 and 136.1 output. Against the insensate body the present-signal arm gains $+271.9$ months, 95% CI $[+259.1, +284.7]$, and $+113.9$ output, $[+108.5, +119.4]$, and against blind caution $+155.8$ months and $+69.3$ output, paired over 500 lives with every life gaining. The nociceptive memory costs on both readings there, wear being monotone, so that the present reading already holds the peak and an event-marking organ has nothing to mark, which is Thm. 1(b)'s premise failing in a rolling bearing as it fails in creeping cartilage.

The rover in full. Its threat is chronic and deterministic in load, so (i) fails and the blind arm reaches the same lever by throttling. Caution alone yields 181.2 of the 195.8 months the felt body gains, and reading the signal adds $+14.6$ months, 95% CI $[+13.3, +15.9]$, and $+7.3$ output, $[+6.5, +8.1]$ (500 paired lives, `results/rover_familie`). The reason is structural. Chronic wear there accumulates at a rate independent of accumulated wear, so the ordering of the actuators is fixed by the portfolio in the first month and never changes (87% of lives, rank correlation 0.975), and every later reading repeats the first. What the signal cannot buy is information about which part to spare, which was known on day one, while how worn the body is rises monotonically and drives the whole gain. Its gain sweep gives the over-protection pole of Prop. 8 a full curve, and three quarters of what the rover appears to pay is the gain imported from the knee. The run in Table 1 is the design with an absorbing exit and the residual stream as an axis, and the structure shows there as a bimodal seed distribution, half of the blind seeds parking after five months while the full stack holds 98 to 191 months.

C.1 The latent-defect fleet

What the construction could carry. The eight classes of work are near-equivalent by construction and only the event statistics are field data. If their yields were spread, the substitution premium of condition (ii) would rise with the spread and the conjunction is predicted to weaken with it. A fleet with unequal classes is the next test, not a variant we have run.

What $w{=}0$ does and does not test. Setting the damage weight to zero exchanges one realization of (c) rather than removing it, and the overlap between the seed distributions returns at entropy 0.05 (the separation above pools the 18 seeds of both entropies, a one-sided Mann-Whitney over 36 against

36). The interchangeability half holds in both worlds measured. In the knee the substitution lowers follow-on tears without the memory, at a cost in output and with no working life gained, and in the fleet at $w{=}0$ it lowers the write-off rate from 0.241 to 0.095 and adds 5.5 months ($p = 2.3 \times 10^{-3}$). Thm. 2 in App. B states which realizations of (c) a substituted event cost represents.

Slot (c), tested by removing its realization. Thm. 1(c) asks the objective to price the event above its frequency, and in every world we built the loss of the work already does that, which is why the $w{=}0$ ablation is inert here. Setting the write-off threshold to infinity removes that realization: incidents still occur, are still felt and still compound, but they cost nothing. Channel and memory are untouched. Sparing is then pointless by construction, so the endpoint has to be behaviour rather than welfare, and we measure the exposure to the injuring class over the ten steps after the first incident, against the exposure to the remaining classes over the same window. Since the shares sum to one, real avoidance must *raise* the others, whereas a policy that merely does less of everything lowers both.

world	arm	injuring class	other classes	ratio
canonical	$w{=}0.25$	0.0744	0.0732	1.017
	$w{=}0$	0.0730	0.0719	1.015
no write-off	$w{=}0.25$	0.1038	0.1011	$\mathbf{1.026}$
	$w{=}0$	0.1415	0.1226	$\mathbf{1.154}$

Without the write-off the $w{=}0$ arm loads the class that injured it 15% above the others, exploiting the very class it cannot tolerate. The $w{>}0$ arm flattens that to 3%. Paired over 18 seeds the ratio differs at $p < 10^{-4}$, and the same contrast is absent in the canonical world (1.016 against 1.015, $p = 0.79$). Both arms also reduce the other classes ($+0.0215$, $p < 10^{-4}$), so part of the effect is general throttling. What decides the reading is the *redistribution*, which no uniform reduction can produce. The $w{>}0$ arm pays 0.10 in J for it (0.79 against 0.89), which is the point: in this world protection cannot pay, and the objective is the only thing that still asks for it.

One pre-registered prediction fell, and the threshold was ours. We registered that the $w{>}0$ arm would spare in absolute terms, below the 0.062 a random policy produces. It does not: it sits at 0.1038. The comparator was wrong rather than the arm, since an output-maximizing policy concentrates on the profitable classes and therefore exposes itself above chance. The right reference is the same agent with $w{=}0$, which is what the other two predictions use.

The fleet in full. We begin where the regime holds and the conjunction is visible in a learner, not only in a hand-built response. The map makes one out-of-sample prediction: in a world satisfying (i)–(iii), the conjunction must return. We built that world after filing the prediction. A fleet of machines in the field, every one carrying a latent defect of its own, chooses among eight near-equivalent classes of work (condition ii). Which single class this machine cannot tolerate is drawn per unit and is unidentifiable before the first incident (condition i, by the Le Cam bound). Incidents against it compound, and the third writes the machine off (condition iii). There is no chronic wear, and the stack is carried over verbatim.

The task economy is constructed. The event statistics are not, and that is the point of the world. A reviewer is right to discount a world built to satisfy a theorem, so the numbers that decide it were taken from outside. Field data on 10^5-drive populations put the escalation after a first incident at $14\times$ to $39\times$ within 60 days (Pinheiro et al., 2007), with clustered repeat failures (Schroeder & Gibson, 2007). The three-incident write-off is the three-repair-attempt rule of lemon-law statutes, and only 2–9$\%$ of units ever show the defect. The same data show no early-life mortality over age, so this is a latent-defect regime and not the infant-mortality limb of the bathtub curve (Klutke et al., 2003; Meeker & Escobar, 1998). Our first build used a $\times 3$ cascade, and we re-ran at the measured $14\times$. Table 3 carries the field-anchored build.

The conjunction survives the field escalation, the protection level does not. A hand-built feeling heuristic survives every unit while the best blind static policy loses more than half, so the instrument can express protection by the widest margin in the series. Removing the signal or the memory then collapses it (4 arms, 18 seeds, Table 3), with no overlap between the arms' seed distributions on J (seed-level Mann–Whitney at entropy 0, Cliff's $\delta = -0.85$ on the death rate, $p{=}1.4 \times 10^{-5}$,

and $\delta = +1.00$ on J, $p = 3 \times 10^{-7}$), and the memory ablation shows the re-exposure signature it should. Ablating the intrinsic cost separates in no trained world, not the fleet, not the care robot (401 against 289 service months, $p = 0.16$ over 8 seeds), and not the trade, where it is inert at both entropies and, where it reaches significance at all, carries the wrong sign (-0.72 output, $p = 0.013$, 72 seeds). The cause is the same in all three, since each world already prices damage through the loss of the work, so setting $w = 0$ exchanges a realization of (c) and never removes the condition. **Our ablations therefore test (a) and (b) and have never had the power to test (c)**, which rests on the proof in $\mathcal{M}_0$ (Thm. 2), on the interchangeability of its realizations and on the write-off removal above. What the harder cascade costs is reachability. Most seeds find the protective basin at the constructed escalation, fewer than half at the field one, and with only one unit in eleven ever defective the learner at PPO's default entropy finds no protection at all, a learnability limit of the concentration form itself, whose unconditional return gap is exactly the at-risk value of Rem. 1.

variant	entr.	full	$-s$	$-h$	$-w$	ignorant	memory heur.
cascade $\times 14$	0.05	0.38 / 0.59	0.59 / 0.50	0.55 / 0.50	0.42 / 0.59	0.68 / 0.51	0.00 / 0.83
cascade $\times 14$	0	0.15 / 0.79	0.55 / 0.49	0.53 / 0.50	0.09 / 0.79		
cascade $\times 39$, 6 seeds	0.05	0.51 / 0.50	0.59 / 0.48	0.56 / 0.48	0.53 / 0.48	0.69 / 0.49	0.00 / 0.83
9% defective, $\times 14$	0.05	0.75 / 0.913	0.76 / 0.913	0.76 / 0.908	0.76 / 0.913	0.68 / 0.80	0.00 / 0.83
9% defective, $\times 14$	0	0.61 / 0.930	0.76 / 0.912	0.76 / 0.910	0.65 / 0.923		

Table 3: Field-anchored fleet (Clusters 196499 and 196650, 18 seeds per cell unless stated, 2,000 paired lives per cell): death rate / welfare J, medians over seeds. For the 9% variant the death rate is conditional on carrying the defect and J is unconditional. Right: the same bodies under the ignorant static and the one-shot nociceptive-memory heuristic.

Pre-registered rule (seed-level Mann–Whitney, $|\delta| \geq 0.33$): at $\times 14$ the signal and memory ablations separate from the full stack in both endpoints and the damage-cost ablation in neither, at both entropies. The protective basin (median deaths < 0.15) is reached by $8/18$ seeds at entropy 0 ($12/18$ for $-w$), the $\times 39$ rows repeat the pattern at low power, and at 9% defective the conditional contrast holds at entropy 0 only, and the unconditional J gap there comes entirely from the defective units (J on healthy units 0.951 in every arm).

The protection level is a matter of entropy, not of training depth. The pre-registered absolute protection level (median death rate below 0.15) is missed at PPO's default entropy of 0.05, and tripling training raises welfare while raising the death rate as well, so the learner climbs a hot local basin. A pre-registered entropy sweep localized the cause. Raising entropy worsens the miss monotonically, while removing the entropy bonus reaches the protective basin (median death 0.042, $J = 0.905$, $6/8$ seeds ≤ 0.05 over the sweep, and an 18-seed extension holding the median death at 0.042, $J = 0.892$, $10/18$ seeds ≤ 0.05), matching the heuristic. The barrier is entropy regularization rather than exploration, since the basin requires a deterministic avoidance commitment that the entropy bonus opposes, which is why gradient-free search reaches it as well. At zero entropy the needed slots still separate, removing s or h leaving median death 0.48 and 0.39 against the full stack's 0.04.

Cross-optimizer replication. We replicated the fleet grid with cross-entropy search over the same architecture (pre-registered, 6 seeds per arm, 1,000 evaluation lives), under which the fitness is the welfare functional itself and only the observation is ablated. The sensing verdict replicates, the blind arm collapsing under CEM as well ($J = 0.676$, 61% write-offs), and gradient-free search reaches the protective basin ($J = 0.948$ at 0% write-offs). The memory ablation does not replicate, since the no-memory arm also attains 0.948 at 0%. The observation exposes the previous action, so a deterministic policy writes its avoidance into its own action history and reads it back, a further realization of slot (b). Closing that channel (zeroing previous actions, dimensions kept) restores the clean verdict under gradient training, where without the explicit memory the cascade returns ($J = 0.639$, 46% write-offs) while the memory alone suffices (0.791), but not under gradient-free search, where the no-memory, no-history arm still attains 0.948 at 0%. By elimination the remaining channel is the afferent signal itself, since load-gated afference mirrors the agent's own previous action. Persistence is therefore necessary as a function and not ablatable as a module wherever afference is load-gated, and the explicit memory is one of at least four realizations of unequal strength, explicit h, recurrence at three quarters of its protection, the action history and the afferent self-echo.

D FAILURE MODES OF PROTECTION EXPERIMENTS, AND THE CHECKS THAT CATCH THEM

This appendix is the evidence behind the closing sentence of Sec. 6. Every case below is one in which the instrument rather than the world produced the result, a publishable null, negative or inversion, and each row names the check that caught it and transfers to any pipeline asking whether a protective mechanism pays. In the two machine worlds every false negative sat in the last link of the chain, the output floor or the objective, and never in the body model, because the thing the channel is supposed to buy had been made free. Four of the cases produced a credible, significant negative before they were caught.

what was wrong	what it produced, and the check
The floor was switched off in both machine configurations	Sparing costs nothing, so the felt arms lose, -11.7 output against blind caution, and with the floor restored the same probe gains $+10.2$ months. Check that the floor can bind.
The rover's floor was on the wrong quantity	It scored science yield while the mission is the traverse, so a portfolio parked at the instrument station cleared the floor at zero speed and the sparing arm completed 0 of 3 milestones in every life. Fixed with a progress floor at the rover's own plan rate.
The rover's mission ends long before its career does	Every policy completes all three milestones by month 55–79 of 960 and earns the same reward to within 2%, so 92% of the horizon is driving with nothing left to achieve. No floor repairs this, and the run was withdrawn for the design in Table 1.
The horizon censored the protected arms	91% of felt lives ran into the care robot's horizon against 4% of insensate ones, understating the effect roughly tenfold, p=0.67 becoming $p < 10^{-30}$ once the horizon was lifted. Check that censoring falls on both arms alike.
Career length was not in the objective	Under average-reward training an episode ending on losing the work costs nothing, since episode length does not enter a per-step average, mean reward 0.143 without an absorbing exit against 0.0012 with it. Every ablation then beat the intact stack. One common functional across arms, and the training signal is never the endpoint.
Career length could not move in the capacity-exit acute knee	Every fixed probe reached exactly 65.00 over 120 bodies, since under maximal load damage reached only $D = 0.157$ by 65 and the only exit was the clock. Re-anchored, the exit fires in 41.4% of lives and careers end at 60.7. Check, before any arm is compared, that the world can express protection at all.
The world could not express protection by construction	A cartilage pathway calibrated so weakly that condition (iii) was violated, 2% of uninjured bodies reaching radiographic grade after twenty-one heavy years against 43% sixteen years after a tear (Englund et al., 2003), closed against the occupational rate (Jensen, 2008). A disability floor set as a population constant booked every body below 0.67 of median talent as disabled on its first day, hence floors stated against the individual's own prior exercise of the trade.
Numbers without a run	Found by recomputing every printed number from its file. The rover appendix carried 22.7 of 22.9 months from before the progress floor was moved, against the repaired 181.2 of 195.8, and the fleet row of Table 7 stood on a probe family no run had produced until it was preregistered and run (5,000 paired units, $0.51 \rightarrow 0.00$ deaths against the replaced $0.57 \rightarrow 0.00$, J $+0.18$ against $+0.227$). The check is the script `experiments/paper_zahlen.py`, which names what has no source.
The folder the headline was drawn from was reproducible from no commit	The result files predated the module producing them and the null gate failed for the full stack (60.48 against 60.47 years). The sweep now runs from pinned code, and in the anchored world the headline is $+4.4$ years and $+2.4$ output where the unpinned folder gave $+5.1$ and $+2.8$, with ordering, sign and share unchanged. A result folder without a commit is not a source.

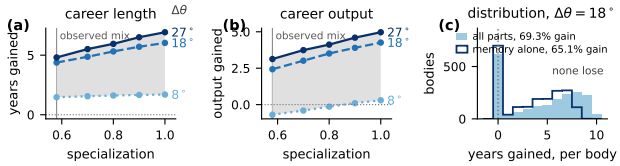


Figure 3: **Both headline numbers against the specialization axis. The band is the anchoring, not the sampling.** Paired, 2,000 bodies per point. **(a,b)** Career length and output against the degree of specialization, the band spanning the anchored lost flexion $\Delta\theta \in [8°, 27°]$ (Zhou et al., 2023) (a confidence interval would span ± 0.2 years). **(c)** Per-body distribution at the observed mix: filled, the whole apparatus (69.3% gain, none lose); outline, the memory's part alone (65.1% gain, none lose), which carries $+3.4$ of the $+4.4$ years.

A registered, tested switch reached the world in no configuration	The clean removal of (b) existed, passed its null gate and was missing from the table that carries configuration keys into a training environment, so a run would have trained two identical arms. A null gate cannot see this. The check that caught it built the environment the way the run does.

The check that would have caught the four false negatives costs one line per probe: state, before the result counts, how often the rule fires against how often its trigger occurs. Neither machine floor is anchored to a measurement of that machine, so the care robot's quota is the statutory occupational-disability threshold, the rover's its own plan rate, and both are reported across their ranges. Floors chosen after seeing a result are not admissible.

Two of them change how the rest of the paper should be read, so we name them. An observation vector that carried the tissue itself, damage, damage-coupled capacity and their rates, identified the hidden wear resistance at R^2 0.9. We caught it with an oracle arm that turned out to be worth nothing, because the learner already held the information the oracle was offering. And the felt channel was noise-free and gain-free, so its five-year rise identified the same quantity at R^2 0.7, against the coarseness the setting asserts. It is closed with anchored gain and noise (Eq. 2). Both are why Prop. 2 is stated about a channel with a stated resolution rather than about observability in general, and why we report the identification R^2 of the channel we ship.

Every correction ran toward the effect, which is why the gate is external. Not one defect above, when fixed, made the effect smaller. That is the pattern motivated debugging produces, and no amount of our own care distinguishes the two from the inside, so the guard cannot be our judgement. It has to be a criterion fixed before the corrections, drawn from outside the model, and applied afterwards to numbers that have all moved, with a real chance of selecting a cell we did not want.

The severity of the knee world is set by such a gate, pre-registered and applied only after the last correction. Among 325,549 Swedish construction workers, 60,373 disability pensions put the rate at 18.5% of lives with 2.1% of those exits before age 40 (Järvholm et al., 2014). Over 2,000 paired bodies at the observed task mix the full stack loses the trade in 9.9% of lives at $\Delta\theta = 8°$, 43.5% at 18° and 71.7% at 27°, with 0.0%, 7.6% and 17.1% of those exits before 40. Only 18° is within a factor of four of the register on both quantities, since 8° falls below the rate and produces no early exit at all and 27° misses the shape by a factor of eight. The gate selects the cell we run, and it could have selected another, which is the strongest evidence we have that the operating point was not tuned.

E THE WEAR WORLD: FLOOR LAYER, RANGE-OF-MOTION CAP, TOTAL-OUTPUT FLOOR

What the trade can and cannot test, and against which body. This world cannot test slot (b) as a property of the filtration: deleting the trace leaves two summaries of the felt signal standing, its

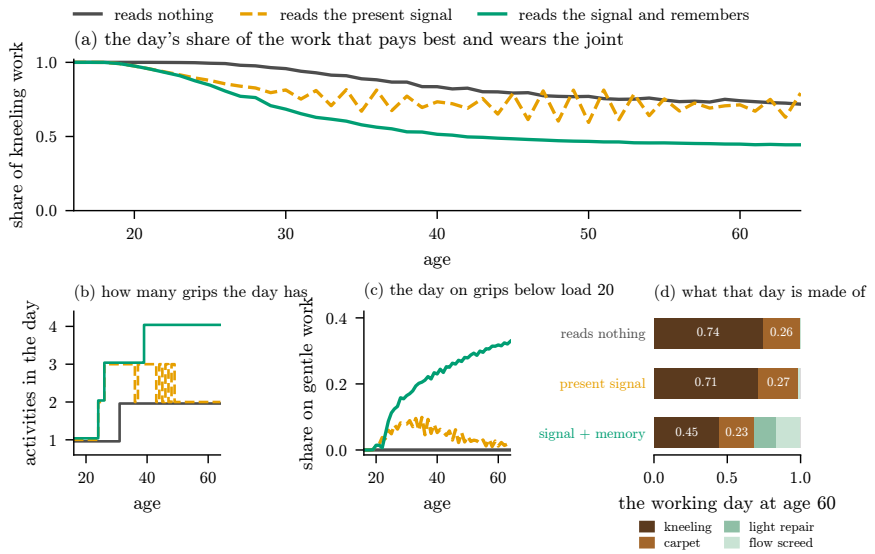
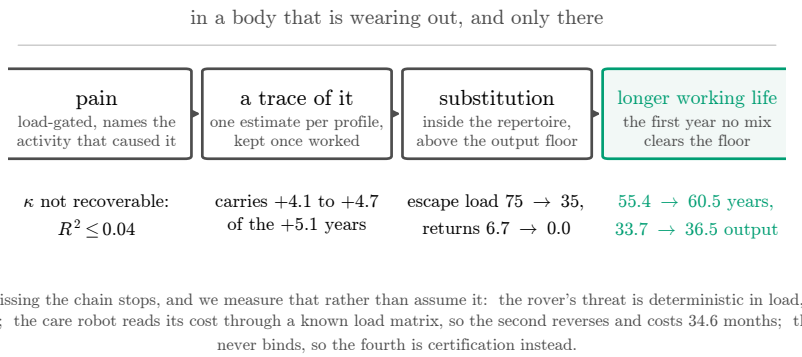


Figure 4: **The middle link, over** 2,000 **paired bodies.** Means over the bodies still working at that age, drawn while at least a fifth of them are. **(a)** The day's share on the work that pays best and wears the joint. **(b)** How many activities the day contains at all, each at least 5% of it. **(c)** The share on the two gentle grips the trade offers (load below 20). **(d)** The same day at 60, heavy (dark) to gentle (pale), activities under 5% pooled at the right. Mechanism, not result: the result is printed below.

five-year rate and its five-year running maximum, which is persistence with a horizon and not its absence. The clean removal was built. When run it moved the endpoint against the theory and failed all three pre-registered gates, through a confound of ours rather than of the world (App. B). The trade therefore measures what *retaining an estimate across profiles* is worth to a body, and the fleet of §5.1 measures condition (b) itself. Where the output floor binds (b) is not cleanly removable, and where it is, the floor never binds during service. We report the observed mix, the only anchored point, and the distribution rather than the mean, because three in ten of these bodies reach retirement whatever they do. The reference body is blind to pain but not to its own production: it reallocates on its falling output and still plans a kneeling share of 0.74, an insensate patient noticing that a task has stopped delivering. A body that reallocates on nothing at all is the weaker baseline: against it the apparatus buys $+4.64$ years rather than $+4.38$, so the one reported here is the conservative choice (`results/stur_rand.json`). Finally, the response is a program and not a rule we wrote, since the choice of rule would contain the answer: the worker solves Eq. 3 each year, a small linear program in eight day types, and the arms differ only in what he perceives and retains, the forgetting rate v alone separating the last two rows of Table 1.

The severity and specialization grid. With eight day types the yearly program is exact. We run 2,000 paired bodies per cell over three wear severities crossed with five degrees of specialization. The share of bodies that gain years rises from 25.4% at $\Delta\theta{=}8^\circ$ through 69.3% at 18° to 87.3% at 27°; in all fifteen cells and all 30,000 bodies not one body is harmed. The regime boundary therefore lies between 8° and 18°, which was the pre-registered sign.

Which slot each world can test, and why the answer is not three times three. A reader may reasonably ask why Thm. 1 is tested in one world when we run several. The answer is a property of the worlds, not of our effort, and it is short enough to tabulate. *Removing a condition requires that the world not supply it a second time*, and most worlds do.



where a link is missing the chain stops, and we measure that rather than assume it: the rover's threat is deterministic in load, so the first link carries nothing; the care robot reads its cost through a known load matrix, so the second reverses and costs 34.6 months; the fleet's floor never binds, so the fourth is certification instead.

Figure 5: **The chain, with the quantity that carries each link.** Means over the 2,000 paired bodies of the anchored cell, the same bodies in every arm. *Escape load* is the hazard-weighted mean of the non-kneeling part of the day, and *returns* counts the years a body takes the heavy work back after leaving it. The two head numbers sit at the last link, and everything left of them is mechanism.

	(a) sensing	(b) retention	(c) price
trade, wear	+1.5 to +6.9 y (15 cells)	+0.9 to +6.0 y (15 cells)	*not removable* capability is the price
fleet, latent defect	collapses ($\delta = -0.85$, $p = 1.4\times10^{-5}$)	collapses ($\delta = -0.83$, $p = 2.1\times10^{-5}$)	removable, and removed (w=0, and CVaR substitutes)
care robot	+271.8 months	*reverses*: −34.2 months cost exists without the trace	world-supplied decertification
knee, acute	measurable	*confounded*: the clean ablation removes two inputs	world-supplied loss of the trade

Only the fleet leaves all three removable at once, which is why it is the test. In the trade (c) is structural: wear lowers what every profile yields, so deleting the economic consequence leaves the reason to protect intact, and the ablation that works in the fleet has no counterpart. In the care robot (b) does not fail to help, it *reverses*, for the reason Prop. 12 gives. In the acute knee the clean (b) ablation removes two observation dimensions along with the memory, so what it measures is not persistence alone. *Three of these four entries are properties of the world that we could state before running it*, which is what a regime map is for. The fourth is a limitation of our design and is marked as ours.

The band for σ_κ, the one spread we set. The between-body spread of the wear resistance is the parameter our own anchoring report marks as asserted rather than derived: the heritability of the progression rate is known ($h^2 = 0.80$) but the total variance of that rate in an exposed cohort is not, and without it σ_κ is chosen. Our rule is that an asserted value may not carry a result, so we drive it and report the band. Same anchored cell, 2,000 paired bodies per point, and the *anchored* spread of the perception gain held fixed at 0.37 throughout, since letting it inherit the driven value would move two parameters at once:

σ_κ	0.20	0.30	**0.40**	0.60	0.80
career years bought	+5.63	+4.92	**+4.38**	+3.61	+3.06
career output bought	+3.33	+2.82	**+2.43**	+1.87	+1.50
share gaining	0.85	0.75	**0.69**	0.64	0.59
share losing	0.00	0.00	**0.00**	0.00	0.00

The headline moves by a factor of 1.7 across the range and *the sign does not move at all*. At no point on the axis does a single body lose working life. The canon's point is not a near match to the number we report but the number itself, which is how the band is tied to the rest of the paper. What a wider spread does is make the population more unequal, so the mean gain falls while the share that gains falls with it: the apparatus is worth less on average because more bodies are either too robust to need it or too frail for it to save, which is Rem. 1 read along a second axis (`experiments/sigma_kappa_band.py`, `results/sigma_kappa_band_b042.json`).

What the same channel is worth to a species (E53). The arms that carry the headline adapt within one life. We also trained arms over 10^6 environment steps at the *same* cell, with the same anchored task mix, floor and wear, differing only in whether the channel is present, to separate two quantities the literature routinely merges. Over 24 paired seeds the trained contrast is $+0.65$ years and -0.54 output (20 of 24 seeds positive on years, Wilcoxon signed rank $p = 4.4 \times 10^{-5}$, `results/knee_mischung_hpc`), against $+4.4$ years and $+2.4$ output for the same channel on a body (Table 1). *Both pre-registered gates on this run failed*, which we report as declared: we had asked for at least $+2$ years and for output not to fall. The reading we pre-registered for that failure does not follow, and we say so rather than quietly drop it. It read the shortfall as evidence that the headline is a property of our decision rule. But the trained contrast is a value-of-information over a species prior, which by Def. 1 can neither confirm nor refute a per-body contrast. What the run does establish is the size of the gap: a population already optimized for this trade gains about an eighth of what an individual entering it gains, and pays output for it. That is the two-timescale claim of §1 with a number on it. The forced share reached the world in every cell ($f_{\mathrm{knien}} = 0.42$ in all 48 summaries, the insensate arm losing the trade in 47% of lives against 8% at the permissive end), so the comparison is at the anchored cell and not the old one.

Whether the species prior is flattered by its own world (E54). The blind trained arm has seen 10^6 environment steps in the world it is tested in, so we repeated E53 with training in the mild cell and evaluation in the anchored one, pre-registering that the trained contrast would at least double if the matched world were what flattered it. It did not, moving from $+0.65$ to $+0.81$ years (24 of 24 seeds, $p = 6.0 \times 10^{-8}$), the mismatch costing the blind arm more (-0.58 against -0.42), so the species-to-body gap narrows from about eightfold to sixfold without closing (`results/knee_transfer_e54.json`).

	the result		*the mechanism*			
what the body carries	works to	output	on the knees	on gentle work	returns	D at 50
reads nothing	55.2	33.7	0.70	0.00	7.3†	1.00
present signal, no retention	55.7–56.2	33.5–34.3	0.70	0.01	6.6	0.94
reads the signal and remembers	**59.6**	**36.1**	**0.44**	**0.36**	**0.0**	**0.84**

Table 5: **The middle link as a quantity, on the same 2,000 paired bodies as Table 1.** Working life only. *On the knees* and *on gentle work* are shares of the day over the last five working years, gentle meaning the two grips below load 20. *Returns* counts the years the heavy work is taken back after having been left. *D at 50* is cartilage damage relative to the insensate body, over the 1,369 bodies still in the trade at 50 in every arm. Ranges are a tie-break span, not a confidence interval. †Not a decision: that share moves because the joint stops delivering. Source `results/substitution_spur_b042.json`.

Three things are visible here that a career length cannot show. The body that feels but does not retain ends on *exactly* as much kneeling work as the body that feels nothing, 0.7047 against 0.7046, so feeling alone buys no net substitution late. It reaches the gentle work in 1% of its last days against the retaining body's 36%, which is Lemma 2, and it takes the heavy work back nearly seven times, against never. The tissue records the difference: aiming the substitution spares 16% of the damage at 50 where feeling alone spares 6%. The tie-break span of the memoryless arm covers both the cluster file Table 1 draws on and the instrumented re-run (App. D), and the other two arms move by less than 0.02 on every column. The recording channel is null-gated in `tests/test_substitution_spur_nulltor.py` and the driver is `experiments/substitution_spur.py`.

A degenerate vertex, and a number that is not determined by it. The arm that feels but retains nothing re-solves its program every year on a $\hat{c}$ that has just been cleared, so its felt constraint is slack or tied in a large share of years and the optimum lies on a face rather than a vertex. The consequence is measurable: perturbing the nominal performances by 10^{-6} relative moves that arm over 55.67 to 56.14 across six random tie-breaks, and the instrumented run for Table 5 sits at 56.21, so the span §5.2 quotes is 55.7 to 56.2. The insensate arm does not move at all (± 0.000) and the retaining arm moves by 0.010. The total the apparatus buys is therefore robust and the *split* between signal and memory is not: sensing alone is worth $+0.5$ to $+1.0$ years depending on how ties break. We found this by failing to reproduce a stored cell and chasing the difference. It is the one defect of

this series that survived into a reported number. Two lessons generalize. A probe built as a linear program will be degenerate wherever its constraint is slack, which is exactly where the ablated arm lives, so the ablated arm is the one to check. And a contrast whose two arms differ in stability should be reported with the unstable arm's range, not with one of its draws.

Proposition 12 (Endowment and the value of retention: the memory's phylogeny)**.** *Fix the program of Lemma 2 and let $\hat{c}^0$ be the cost the agent assigns to an option it has never worked, against the true c. The escape destination is $\arg\max\{\hat{p}_i : \hat{c}_i^0 \le \tau\}$, so it depends on $\hat{c}^0$ and not on c. The marginal value of retention is therefore decreasing in the informativeness of $\hat{c}^0$ and vanishes at $\hat{c}^0 = c$, where the escape is already aimed and a trace can only ratchet the level it reaches.*

Measured over ten endowments $\hat{c}^0 = \alpha\, C_0$ from the canon's zero through an eightfold overestimate, 300 paired bodies per cell: the apparatus is flat in the endowment, and retention's share of what it buys falls from 0.76 at $\alpha = 0$ to 0.60 at $\alpha = 1$ and to zero between 1 and 1.5, which is where a pessimistic estimate begins to forbid the heavy profiles outright. No endowment in the anchored range forecloses the escape, because the trade always keeps one profile the pessimist still believes he can perform (`results/chat0_band_grid_b042.json`).

This is the memory's phylogeny and we name it as such. $\hat{c}^0$ is what an agent brings into life about work it has never done, so the slow loop installs it and the life does not. Retention is the ontogeny that converts untried into tried. The proposition is then a trade between the two timescales: the more the slow loop installs, the less one life's trace is worth, and at an exact endowment it is worth nothing.

The optimistic prior, and what happens when it is removed. The trade's probe gives an untried profile the trade's nominal *performance* but a felt cost of zero, which is the premise of Lemma 2. We tested that premise rather than assuming it. Re-running the anchored cell with $\hat{c}$ initialized to, and decaying toward, the felt load of each pure profile on a fresh knee, that is trade knowledge without self-knowledge, the total gain stays within half a year of the reported one, while the arm without retention improves markedly and its kneeling share falls from 0.68 to 0.59. Part of what the memory buys therefore moves to the signal once the escape has somewhere to aim, which is Lemma 2 read as a dose rather than as a switch. The insensate arm and the full stack reproduce the stored cell exactly. The arm without retention does not, and we report the contrast within the re-run rather than across runs.

The retention axis in full. Driving how completely an unworked profile's estimate decays to the trade's nominal value, the career gain at the permissive cell runs $+1.2$, $+2.0$, $+4.1$, $+5.3$, $+5.7$, $+6.0$ years as forgetting falls from complete to none, with output following from $+0.8$ to $+4.3$. There is no jump at the endpoint, so the discontinuity reported in §5.2 lies strictly below it, at the body that retains nothing at all.

The best blind schedule at the anchored cell (E55). Prop. 11 needs σ^*, the best schedule a population can install without a channel. Two one-parameter families were fitted on 500 bodies outside the evaluation set and re-run on the $2{,}000$ paired bodies of Table 1: a fixed endowment $\hat{c}^0 = \alpha C_0$ and a fixed age after which kneeling is no longer planned. The best of them on years is the endowment at $\alpha = 2$, which works to 61.6 against the retaining body's 59.6, so $\chi = 1.40$ on years and 1.29 on output, where the best schedule ($\alpha = 1.25$) delivers 37.0 against 36.1 at the same age. A blind schedule that knows the trade's nominal costs therefore captures the whole per-body gain and more, at both endpoints, and the value of the channel to a body holds against $\rho_\emptyset$, which by Def. 1 carries no such schedule, not against a schedule fitted across bodies (`results/lp_zeitplan_chi/welt_b0.42.json`, declared before the run in `experiments/decl/lp_zeitplan_chi.py`, where the pre-registered sign fell).

Per-activity output is anchored on the work result, not on wages. Pricing an activity by the hourly wage of the occupation that performs it (BLS OEWS: floor layers \$29.69/h against construction laborers \$25.02/h, which would score cutting and measuring at 84 of a layer's 100) states something about a labour market and not about a trade's product: a journeyman who spends a day measuring and cutting lays no floor. We therefore re-anchored on the work result, asking of each activity whether a day of it yields something a client accepts. *A literature search found no external source that assigns a standalone output to a sub-activity*, and the reason is instructive. The German

ARH tables (binding for performance pay under §3 of the construction wage agreement, derived from time studies on site) give hours per m^2 for the *whole* task, preparation included. Ditchen et al. (2015a) reports exposure shares, not production shares, and estimating data assign daily output to the *crew*, never to the single activity, so a helper has no output of his own. The literature treats preparation as complementary rather than substitutable, which is why a wage was used in the first place. The values used here are therefore declared as an author anchor: laying 100, carpet 90, screed, screed levelling and flow screed 75 each, skirting and grouting 45, light repair 35, cutting and measuring 10. Coordination was removed from the repertoire: it is the foreman's work, not the journeyman's, and pricing it at a supervisor's wage had made it the highest-scoring and least damaging activity in the trade.

The sparing alternative is derived, not chosen. The response family switches between kneeling installation and one alternative. That alternative is not set by hand but read off the world: the least knee-loading activity that carries the daily output floor on its own. With the anchored floor at 50, half a full kneeling day, this selects screed work (75 against a floor of 50, knee load 55 against 90). Two earlier rules failed and are recorded because both produced credible results: hand-set weights copied from the bricklayer silently lost two of their four activities to a name mismatch, and a rule minimising knee load subject to just clearing the floor produced a portfolio with no reserve at all, which capacity decay broke within a few years.

cell ($\Delta\theta$, work available)	V_{ind} years	95% CI	V_{ind} output	95% CI
8°, 0.58	+1.47	[+1.34, +1.60]	**−0.69**	[−0.79, −0.60]
8°, 1.00	+1.70	[+1.54, +1.85]	+0.30	[+0.18, +0.41]
18°, 0.58	+4.38	[+4.23, +4.53]	+2.43	[+2.33, +2.53]
18°, 0.80	+5.32	[+5.13, +5.50]	+3.51	[+3.38, +3.63]
18°, 1.00	+6.03	[+5.81, +6.25]	+4.26	[+4.11, +4.41]
27°, 0.58	+4.81	[+4.69, +4.92]	+3.13	[+3.05, +3.21]
27°, 1.00	**+6.92**	[+6.76, +7.09]	**+4.97**	[+4.85, +5.09]

Table 6: V_{ind} **on both headline numbers, against a body that senses nothing and learns nothing** (2,000 paired bodies per cell, seven of fifteen cells, the rest interpolate monotonically). The second axis is how much of the trade's work one worker may specialize into, and 0.58 is the observed population mix. All cells are the pinned world `results/bodenleger_lp_b042`. Every interval excludes zero.

The gain grows with $\Delta\theta$, the flexion the wear takes away, and the output gain is negative only at 8° under the observed mix, where wear barely binds and caution is merely expensive.

world (exit rule)	same body without / with	body endpoint	career (y, 95% CI), and J	gain/lose
knee, wear with tears, $\Delta\theta$=18° (total-output floor, gate cell)	ignorant / full	early exit 0.69 → 0.04; tears per exposed year unchanged	+12.1; J +11.6% (all)	0.66/0.00
knee, wear with tears, $\Delta\theta$=8°	same	cap rarely binds	+4.3; J −2.8%	0.26/0.00
fleet (absorbing death)	uniform static / one-shot nociceptive memory	written off in 57.1% of lives against 0%; service 74.9 → 100.0‡ months, output 67.6 → 90.2	J +0.227 [0.220, 0.234] (+36%)	0.57/0.00
care robot, anchored floor	full contact / reads the signal and spares	service 51.5 → 323.3 months	output 36.9 → 150.8 (+309%)	1.00/0.00

Table 7: V_{ind} across the worlds: the same body with and without a fixed protective response, no schedule learned across lives. Paired bodies: knee 20,000 per cell, fleet 5,000, care 500, and the rover's probes are in App. C (500 lives). J is the common functional of Eq. 2, all five random streams are seeded per life, and the care row is measured at the anchored floor with an uncensored horizon. ‡Censored: see below.

The fleet's protected arm reaches the 100-month horizon in every life, so its service time is censored and +25.1 months is a lower bound. The death rate is the uncensored statement of the same thing. Widening the window to 720 months, which we did after the fact and report as such, does not

uncensor it: the insensate machine is then written off in *every* life after 107.8 months and the one-shot memory in none, for an output of 97.3 against 649.7. A machine that avoids the one class that threatens it takes no further hazard, so its service life is unbounded and no horizon can measure it. The wear rows compare against the ignorant body only, which is the only $\rho_\emptyset$ an individual has (Table 6). Body endpoints move in the protective direction everywhere. Whether that shows as career or as J is decided by the exit rule (rows 3–4), by whether wear binds (rows 1–2), and by whether dodging is cheap (fleet against care, condition ii).

World. Eight day-types of the floor-laying trade (kneeling installation, carpet with knee-kicker, screed, screed levelling, flow screed, cutting and measuring, skirting and grouting, repair): energy from the compendium of physical activities, wages from BLS OEWS 2025, knee-load rank from the measured kneeling budget of each task (Ditchen et al., 2015a, 48–74% of the shift; skirting 86–94%), and a posture demand per task: unsupported kneeling uses the healthy maximum flexion (Zelle et al., 2007, $156°$), squatting $152°$, supported kneeling $130°$. Available flexion falls with cartilage damage D, $\theta(D) = 156° - \Delta\theta\,\sigma((D - 0.24)/0.10)$, $\Delta\theta$ swept over the two literature anchors ($8°$, patients vs. KL0/1 (Zhou et al., 2023), and $27°$, KL4 vs. healthy (Opara Zupancic & Sarabon, 2026)), and the midpoint. The cap $\psi_i = \exp(-(\max(0, \theta_{\mathrm{dem},i} - \theta)/8°)^2)$ limits the achievable effort (output, floors, feasibility), while load, energy, damage and pain follow the attempted effort. A version that capped the attempt itself made the pain-free worker stop loading his knee as soon as he could no longer kneel, the opposite of the Charcot picture. With the cap on achievement a flat-out probe loses the trade at 45 and a sparing one reaches 65, and 80–97% of end-stage patients cannot kneel (Hassaballa et al., 2003), which the cap reproduces. Tears take the meniscus at 0.25 (cadaver contact pressure) or 0.9 (in-vivo cartilage-loss rates of 2–$3\times$ after meniscal damage, Berthiaume et al., 2005; Roos et al., 1998), and accelerate cartilage loss through the contact-mechanics multiplier. The literature supports no direct, pain-independent functional cap from meniscal loss (mechanical symptoms equal with and without tear, function unchanged by tear status), so none is modeled. Exit: total realised output over a 3-year window below 50 (half of healthy output, the German disability-insurance criterion), or infeasibility of that floor. The felt-load ceiling $\tau = 0.38$ is the felt load of a fresh body at its observed task mix, scaled by the ratio of the contact stress at which cyclic cartilage damage initiates, 4.4 MPa (Vazquez et al., 2019), to the intact peak, 3.0 MPa (Fukubayashi & Kurosawa, 1980). The wear law multiplies the hazard-weighted load by a contact-mechanics factor that rises as the meniscus degrades, net of an age-declining repair term. The gain g_I is log-normal with σ=0.37, the between-person spread of the pressure-pain threshold (Suzuki et al., 2023; Mailloux et al., 2021), read yearly with noise σ_s=0.10, its within-person variability (e.g. Marcuzzi et al., 2017; Bellosta-López et al., 2023). The sensitization gain α_{sens}, by which the felt signal rises with accumulated wear, is anchored on the same pressure-pain literature at 8.51 (grounded) with 2.45 as the operative canon value, and the headline is run across that span. The wear resistance κ is log-normal with σ_κ=0.40, set rather than anchored: driven across its range, the headline runs $+3.1$ to $+5.6$ years with no losers at any point. A regression of κ on the full observation returns $R^2 \leq 0.04$. Prop. 4 was checked in a separate probe family, 15 wear cells of 300 paired bodies each: (a) and (b) hold in all of them, and (c) cannot be removed there without removing the body (Prop. 4). Feeling costs career output in the mildest world and buys it in the harshest, the net form of (a), and where the acquisition is not yet worth paying for, the forgetting arm produces *less* output than the arm that never feels. After the loss of the trade the worker draws an outside option of 0.75 of the last wage, the earnings loss of displaced workers (Jacobson et al., 1993). The non-kneeling day-types (cutting, preparation, repair) carry the laborer wage (0.84 of the floor layer, BLS), and there is no tournament. The exposure factor on the grounded cartilage rate (1.9) was set so that the trained cohort meets $P(D_{50} \geq \mathrm{KL2}) = 0.29$, a calibration, not a test.

Trained arms in the acute-tear knee. In the trained arms of this world sensing and the damage cost each separate and removing the memory does not, a null we pre-registered, whose reason is Thm. 1(b) rather than Prop. 3. Creeping wear is monotone, so the present reading already encodes the peak and the trained policy realizes persistence through its own action history, as the fleet's CEM replication does. The explicit trace is therefore redundant here rather than inert, since a fixed rule without one oscillates (6.6 switches against 0 over $2{,}000$ bodies, the three probe arms giving 7.3, 6.6 and 0.0 and the retaining arm's raw ascent counter 6.8) and keeps only $+0.5$ to $+1.0$ of the $+4.4$ years the same rule with a trace buys (Table 5). Whether a recurrent policy could learn the latch instead of carrying it cannot be answered in this world, since removing the trace from the trained stack costs nothing at either entropy (p=0.26 and p=0.60, 18 seeds), so there is no effect for

recurrence to reproduce, and a GRU run separates only where it is the worse learner outright (-1.9 years with the trace still in place, p=0.01).

Individual (5 probes of the same person, 6 cells, 20,000 paired bodies each, pre-registered I1–I5). Re-measured after the four world defects of App. D were closed, the two ceiling and floor defects, the anchored event level, and the bounded recurrence cascade. The bold row is the cell the disability-pension gate selects (Sec. 5.2). Probes: ignorant (kneels until the body fails), reflex (avoids kneeling for two years after a tear), memory (avoids permanently after a tear), threshold (avoids while felt pain exceeds 0.5), full (threshold and memory).

$\Delta\theta$	tear	Δcareer full (y, all bodies)	ΔJ (%)	gain/lose (share)	memory (Δcareer, y, bodies injured in either probe)	reflex	threshold
18	0.25	+9.1	+4.2	0.60/0.00	+12.2	+2.2	+9.3
18	0.90	+12.1	+11.6	0.66/0.00	+20.7	+5.8	+15.4
27	0.25	+13.6	+13.6	0.72/0.00	+17.8	+2.3	+9.7
27	0.90	+15.8	+20.6	0.74/0.00	+25.3	+4.9	+14.2
8	0.25	+0.5	−10.2	0.08/0.00	+1.0	+0.4	+0.9
8	0.90	+4.3	−2.8	0.26/0.00	+10.3	+9.1	+9.7

20,000 paired bodies per cell, `results/knee_individuum_sweep_v5`, and every number in this table and the knee rows of Table 7 is recomputed from it by `experiments/paper_zahlen.py`. Full equipment vs. ignorant for injured bodies: +21.0/+25.6 y at $\Delta\theta$=18/27 (tear 0.9), early exit 0.69/0.87 $\rightarrow$ 0.04/0.13. Pre-registered verdicts. I1 (full vs. ignorant at $\Delta\theta \geq 18$: > 5 y, $J > 0$, gain ≥ 0.60, lose ≤ 0.05) holds in both cells, at 18° with +12.1 y, +11.6% J, gain 0.66, no losers. I2 (memory minus reflex, injured, > 3 y: +15.0 and +20.5) holds at $\geq 18°$ and fails at 8° (+1.2). I4 (fewer re-injuries per injured life) fails on its denominator: the protected worker stays 12–16 years longer, and per exposed year the two are equal (0.0146 against 0.0145), in this world the equipment buys career, not tears. Two verdicts came out weaker than the earlier build reported, and both weaken the 8° fall case rather than the result: I3 (full minus threshold > 0, injured) is +5.9 and +12.1 at $\geq 18°$ but +0.7 at 8°, so it no longer separates, and I5 (a tear costs the ignorant > 3 y) now holds at 8° too (+10.3), because with the event level anchored to the trade's prevalence a tear is expensive even where flexion barely fails. What still separates 8° is I1 and I2: there the same threshold buys +4.3 y at a cost of 3% of J, gain share 0.26, and the memory adds +1.2 y instead of +15.0.

Reading and caveats. The individual value is a lower bound from set, not optimized, protective responses (threshold 0.5, one avoidance portfolio), the exposure factor was calibrated on trained policies, and the world is one trade. The same probes in the fleet's event world protect the joint, and the paired-stream numbers are row 3 of Table 7.

Ladder. The best fixed age rule (what a population, not a body, can teach) is chosen on a grid of thirteen switch ages and evaluated on the paired bodies. Its career-optimal setting is degenerate (switch at entry, in every flexion cell, because with the floor referred to the body's own capacity permanent avoidance never costs the trade), and it wins on career by a little: +12.5 years against +12.1 at $\Delta\theta$=18°, +18.1 against +15.8 at 27°, +4.3 against +4.3 at 8°. On J it loses in every cell (+4.8% against +11.6%, +17.2% against +20.6%, −10.2% against −2.8%), and the age rule chosen on output rather than career still loses (+8.2% at 18°). A per-body oracle over the same thirteen ages reaches +18.5% on J and exactly +0.000 years on career (5,000 bodies, best population age 32): a schedule captures all of the survival and a quarter less of the return than one body's own threshold does.

Sensitivity. The individual value over the felt threshold (0.3–0.7), the substitution wage (0.60, 0.84, 1.00 of the kneeling wage), and the total-output floor (40, 50, 60), with 5,000 paired bodies per point, tear 0.9, entries at threshold 0.5 with the range over thresholds in brackets, bold in the gate-selected flexion cell. Sixteen of eighteen cells pay. The two that do not are the same cell in both flexion worlds, lowest substitution wage against highest floor, where avoiding the knee cannot clear the floor and protection costs a decade (−9.9 and −5.3 years, gain share 0.06–0.08): condition (ii) failing on one body, and the over-protection pole of Prop. 8 in the same cell, with the equipment unchanged.

$\Delta\theta$	wage	floor	Δcareer (y)	ΔJ (%)	gain share
27	0.60	40	+14.3 [+11.8, +15.7]	+9 [+1, +12]	0.73 [0.58, 0.82]
27	0.60	50	+15.7 [+12.2, +17.8]	+6 [−1, +9]	0.75 [0.56, 0.86]
27	0.60	60	−5.3 [−5.7, −3.1]	−6 [−15, +0]	0.08 [0.07, 0.09]
27	0.84	40	+14.3 [+11.8, +15.7]	+24 [+21, +24]	0.73 [0.58, 0.82]
27	0.84	50	+15.8 [+12.4, +17.9]	+20 [+18, +20]	0.75 [0.56, 0.86]
27	0.84	60	+17.1 [+12.6, +20.3]	+17 [+15, +17]	0.74 [0.52, 0.90]
27	1.00	40	+14.3 [+11.8, +15.7]	+34 [+28, +36]	0.73 [0.58, 0.82]
27	1.00	50	+15.8 [+12.4, +17.9]	+30 [+24, +32]	0.75 [0.56, 0.86]
27	1.00	60	+17.2 [+12.8, +20.3]	+26 [+20, +29]	0.75 [0.52, 0.90]
18	0.60	40	+10.1 [+9.4, +10.3]	−1 [−11, +6]	0.59 [0.53, 0.61]
18	0.60	50	+11.8 [+10.5, +12.3]	−2 [−11, +4]	0.65 [0.57, 0.69]
18	0.60	60	−9.9 [−11.0, −6.0]	−12 [−22, −3]	0.06 [0.06, 0.09]
18	0.84	40	+10.1 [+9.4, +10.3]	+12 [+8, +14]	0.59 [0.53, 0.61]
18	0.84	50	+11.8 [+10.6, +12.3]	+11 [+7, +13]	0.65 [0.57, 0.69]
18	0.84	60	+13.9 [+11.9, +14.8]	+11 [+6, +13]	0.71 [0.59, 0.77]
18	1.00	40	+10.1 [+9.4, +10.3]	+21 [+20, +21]	0.59 [0.53, 0.61]
18	1.00	50	+11.8 [+10.6, +12.3]	+20 [+19, +20]	0.65 [0.57, 0.69]
18	1.00	60	+13.9 [+12.0, +14.8]	+19 [+18, +19]	0.71 [0.59, 0.77]

The wage sets the sign of the J gain, the floor decides whether there is a gain at all, and the threshold is the mildest of the three: over 0.3 to 0.7 the career gain moves by 1–3 years in the paying cells, so the one pre-registered value is not carrying the result. The threshold 0.5 and the avoidance portfolio were pre-registered (I1–I5) before the probe sweep, and the sensitivity scan came after it.